



1/53

APPROVED BY DRAFTSMAN	O.G. FIG.	SUBCLASS
	CLASS	

TCGCCCCACGCGTCCGCACCGCCGCCAGGCAAGGCCGCCCTGCCTTGGG

↑SEQ.ID NO: 1

SEQ.ID NO: 2 → M A G G R G A P 8

CGCAGCGCTGCC ATG GCT GGG GGC CGT GGG GCC CCC 24

SEQ.ID NO: 3 ↑→

G R G R D E P P E S Y 19

GGG CGC GGC CGG GAC GAG CCT CCG GAG AGC TAC 57

P Q R Q D H E L Q A L 30

CCG CAA CGA CAG GAC CAC GAG CTA CAG GCC CTG 90

E A I Y G A D F Q D L 41

GAG GCC ATC TAC GGC GCG GAC TTC CAA GAC CTG 123

R P D A C G P V K E P 52

CGG CCG GAC GCT TGC GGA CCG GTC AAA GAG CCC 156

P E I N L V L Y P Q G 63

CCT GAA ATC AAT TTA GTT TTG TAC CCT CAA GGC 189

L T G E E V Y V K V D 74

CTA ACT GGT GAA GAA GTA TAT GTA AAA GTG GAT 222

L R V K C P P T Y P D 85

TTG AGG GTT AAA TGC CCA CCT ACC TAT CCA GAT 255

V V P E I E L K N A K 96

GTA GTT CCT GAA ATA GAG TTA AAA AAT GCC AAA 288

G L S N E S V N L L K 107

GGT CTA TCA AAT GAA AGT GTC AAT TTG TTA AAA 321

S R L E E L A K K H C 118

TCT CGC CTA GAA GAA CTG GCC AAG AAA CAC TGT 354

G E V M I F E L A Y H 129

GGG GAG GTG ATG ATC TTT GAA CTG GCT TAC CAC 387

V Q S F L S E H N K P 140

GTG CAG TCA TTT CTC AGC GAG CAT AAC AAG CCC 420

FIG. 1A



P	P	K	S	F	H	E	E	M	L	E	151
CCT	CCC	AAG	TCT	TTT	CAT	GAA	GAA	ATG	CTG	GAA	453
R	R	A	Q	E	E	Q	Q	R	L	L	162
AGG	CGG	GCT	CAG	GAG	GAG	CAG	CAG	AGG	CTG	TTG	486
E	A	K	R	K	E	E	Q	E	Q	R	173
GAG	GCC	AAG	CGG	AAA	GAA	GAG	CAG	GAG	CAA	CGT	519
E	I	L	H	E	I	Q	R	R	K	E	184
GAA	ATC	CTG	CAT	GAG	ATT	CAG	AGA	AGG	AAA	GAA	552
E	I	K	E	E	K	K	R	K	E	M	195
GAG	ATA	AAA	GAA	GAG	AAA	AAA	AGG	AAA	GAA	ATG	585
A	K	Q	E	R	L	E	I	A	S	L	206
GCT	AAG	CAG	GAA	CGT	TTG	GAA	ATT	GCT	AGT	TTG	618
S	N	Q	D	H	T	S	K	K	D	P	217
TCA	AAC	CAA	GAT	CAT	ACC	TCT	AAG	AAG	GAC	CCA	651
G	G	H	R	T	A	A	I	L	H	G	228
GGA	GGA	CAC	AGA	ACG	GCT	GCC	ATT	CTA	CAT	GGA	684
G	S	P	D	F	V	G	N	G	K	H	239
GGC	TCT	CCT	GAC	TTT	GTA	GGA	AAT	GGT	AAA	CAT	717
R	A	N	S	S	G	R	S	R	R	E	250
CGG	GCA	AAC	TCC	TCA	GGA	AGG	TCT	AGG	CGA	GAA	750
R	Q	Y	S	V	C	N	S	E	D	S	261
CGT	CAG	TAT	TCT	GTA	TGT	AAT	AGT	GAA	GAT	TCT	783
P	G	S	C	E	I	L	Y	F	N	M	272
CCT	GGC	TCT	TGT	GAA	ATT	CTG	TAT	TTC	AAT	ATG	816
G	S	P	D	Q	L	M	V	H	K	G	283
GGG	AGT	CCT	GAT	CAG	CTC	ATG	GTG	CAC	AAA	GGG	849

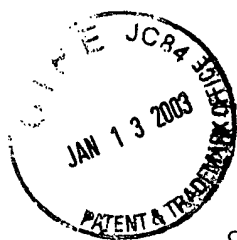


3/53

APPROVED	CLASS	SUBCLASS
BY		
CRAFTSMAN		

K	C	I	G	S	D	E	Q	L	G	K	294
AAA	TGT	ATT	GGC	AGT	GAT	GAA	CAA	CTT	GGA	AAA	882
L	V	Y	N	A	L	E	T	A	T	G	305
TTA	GTC	TAC	AAT	GCT	TTG	GAA	ACA	GCC	ACT	GGT	915
G	F	V	L	L	Y	E	W	V	L	Q	316
GGC	TTT	GTC	TTG	TTG	TAT	GAG	TGG	GTC	CTT	CAG	948
W	Q	K	K	M	G	P	F	L	T	S	327
TGG	CAG	AAA	AAA	ATG	GGT	CCA	TTC	CTT	ACC	AGT	981
Q	E	K	E	K	I	D	K	C	K	K	338
CAA	GAA	AAA	GAG	AAG	ATT	GAT	AAG	TGC	AAA	AAG	1014
Q	I	Q	G	T	E	T	E	F	N	S	349
CAG	ATT	CAA	GGA	ACA	GAA	ACA	GAA	TTC	AAC	TCA	1047
L	V	K	L	S	H	P	N	V	V	R	360
CTG	GTA	AAA	TTG	AGC	CAT	CCA	AAT	GTA	GTA	CGC	1080
Y	L	A	M	N	L	K	E	Q	D	D	371
TAC	CTT	GCA	ATG	AAT	CTC	AAA	GAG	CAA	GAC	GAC	1113
S	I	V	V	D	I	L	V	E	H	I	382
TCC	ATC	GTG	GTG	GAC	ATT	TTA	GTG	GAG	CAC	ATT	1146
S	G	V	S	L	A	A	H	L	S	H	393
AGT	GGG	GTC	TCT	CTT	GCT	GCA	CAC	CTG	AGC	CAC	1179
S	G	P	I	P	V	H	Q	L	R	R	404
TCA	GGC	CCC	ATC	CCT	GTG	CAT	CAG	CTT	CGC	AGG	1212
Y	T	A	Q	L	L	S	G	L	D	Y	415
TAC	ACA	GCT	CAG	CTC	CTG	TCA	GGC	CTT	GAT	TAT	1245
L	H	S	N	S	V	V	H	K	V	L	426
CTG	CAC	AGC	AAT	TCT	GTG	GTG	CAT	AAG	GTC	CTG	1278

FIG. 1C



4/53

CLASS	SUBCLASS
BY	CRAFTSMAN

S	A	S	N	V	L	V	D	A	E	G	437
AGT	GCA	TCT	AAT	GTC	TTG	GTG	GAT	GCA	GAA	GGC	1311
T	V	K	I	T	D	Y	S	I	S	K	448
ACC	GTC	AAG	ATT	ACG	GAC	TAT	AGC	ATT	TCT	AAG	1344
R	L	A	D	I	C	K	E	D	V	F	459
CGC	CTC	GCA	GAC	ATT	TGC	AAG	GAG	GAT	GTG	TTT	1377
E	Q	T	R	V	R	F	S	D	N	A	470
GAG	CAA	ACC	CGA	GTT	CGT	TTT	AGT	GAC	AAT	GCT	1410
L	P	Y	K	T	G	K	K	G	D	V	481
CTG	CCT	TAT	AAA	ACG	GGG	AAG	AAA	GGA	GAT	GTT	1443
W	R	L	G	L	L	L	L	S	L	S	492
TGG	CGT	CTT	GGC	CTT	CTG	CTG	CTG	TCC	CTC	AGC	1476
Q	G	Q	E	C	G	E	Y	P	V	T	503
CAA	GGA	CAG	GAA	TGT	GGA	GAG	TAC	CCT	GTG	ACC	1509
I	P	S	D	L	P	A	D	F	Q	D	514
ATC	CCT	AGT	GAC	TTA	CCA	GCT	GAC	TTT	CAA	GAT	1542
F	L	K	K	C	V	C	L	D	D	K	525
TTT	CTA	AAG	AAA	TGT	GTG	TGC	TTG	GAT	GAC	AAG	1575
E	R	W	S	P	Q	Q	L	L	K	H	536
GAA	AGA	TGG	AGT	CCC	CAG	CAG	TTG	TTG	AAA	CAC	1608
S	F	I	N	P	Q	P	K	M	P	L	547
AGC	TTT	ATA	AAT	CCC	CAG	CCA	AAA	ATG	CCT	CTA	1641
V	E	Q	S	P	E	D	S	G	G	Q	558
GTG	GAA	CAA	AGT	CCT	GAA	GAT	TCT	GGA	GGA	CAA	1674
D	Y	V	E	T	V	I	P	S	N	R	569
GAT	TAT	GTT	GAG	ACT	GTT	ATT	CCT	AGC	AAC	CGG	1707
L	P	S	A	A	F	F	S	E	T	Q	580
CTA	CCC	AGT	GCT	GCC	TTC	TTT	AGT	GAG	ACA	CAG	1740

FIG. 1D



5/53

APPROVED BY DRAFTSMAN	O.G. FIG.	SUBCLASS
	CLASS	

R	Q	F	S	R	Y	F	I	E	F	E	591
AGA	CAG	TTT	TCC	CGA	TAC	TTC	ATT	GAG	TTT	GAA	1773
E	L	Q	L	L	G	K	G	A	F	G	602
GAA	TTA	CAA	CTT	CTT	GGT	AAA	GGA	GCT	TTT	GGA	1806
A	V	I	K	V	Q	N	K	L	D	G	613
GCT	GTC	ATC	AAG	GTG	CAG	AAC	AAG	TTG	GAC	GGC	1839
C	C	Y	A	V	K	R	I	P	I	N	624
TGC	TGC	TAC	GCA	GTG	AAG	CGC	ATC	CCC	ATC	AAC	1872
P	A	S	R	Q	F	R	R	I	K	G	635
CCG	GCC	AGC	CGG	CAG	TTC	CGC	AGG	ATC	AAG	GGC	1905
E	V	T	L	L	S	R	L	H	H	E	646
GAA	GTG	ACA	CTG	CTG	TCA	CGG	CTG	CAC	CAT	GAG	1938
N	I	V	R	Y	Y	N	A	W	I	E	657
AAC	ATT	GTG	CGC	TAC	TAC	AAC	GCC	TGG	ATC	GAG	1971
R	H	E	R	P	A	G	P	G	T	P	668
CGG	CAC	GAG	CGG	CCG	GCG	GGA	CCG	GGG	ACG	CCG	2004
P	P	D	S	G	P	L	A	K	D	D	679
CCC	CCG	GAC	TCC	GGG	CCC	CTG	GCC	AAG	GAT	GAC	2037
R	A	A	R	G	Q	P	A	S	D	T	690
CGA	GCT	GCA	CGC	GGG	CAG	CCG	GCG	AGC	GAC	ACA	2070
D	G	L	D	S	V	E	A	A	A	P	701
GAC	GGC	CTG	GAC	AGC	GTA	GAG	GCC	GCC	GCG	CCG	2103
P	P	I	L	S	S	S	V	E	W	S	712
CCA	CCC	ATC	CTC	AGC	AGC	TCG	GTG	GAG	TGG	AGC	2136
T	S	G	E	R	S	A	S	A	R	F	723
ACT	TCG	GGC	GAG	CGC	TCG	GCC	AGT	GCC	CGT	TTC	2169

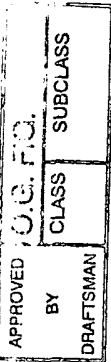
FIG. 1E



6/53

P	A	T	G	P	G	S	S	D	D	E	734
CCC	GCC	ACC	GGC	CCG	GGC	TCC	AGC	GAT	GAC	GAG	2202
D	D	D	E	D	E	H	G	G	V	F	745
GAC	GAC	GAC	GAG	GAC	GAG	CAC	GGT	GGC	GTC	TTC	2235
S	Q	S	F	L	P	A	S	D	S	E	756
TCC	CAG	TCC	TTC	CTG	CCT	GCT	TCA	GAT	TCT	GAA	2268
S	D	I	I	F	D	N	E	D	E	N	767
AGT	GAT	ATT	ATC	TTT	GAC	AAT	GAA	GAT	GAG	AAC	2301
S	K	S	Q	N	Q	D	E	D	C	N	778
AGT	AAA	AGT	CAG	AAT	CAG	GAT	GAA	GAT	TCC	AAT	2334
E	K	N	G	C	H	E	S	E	P	S	789
GAA	AAG	AAT	GGC	TGC	CAT	GAA	AGT	GAG	CCA	TCA	2367
V	T	T	E	A	V	H	Y	L	Y	I	800
GTG	ACG	ACT	GAG	GCT	GTG	CAC	TAC	CTA	TAC	ATC	2400
Q	M	E	Y	C	E	K	S	T	L	R	811
CAG	ATG	GAG	TAC	TGT	GAG	AAG	AGC	ACT	TTA	CGA	2433
D	T	I	D	Q	G	L	Y	R	D	T	822
GAC	ACC	ATT	GAC	CAG	GGA	CTG	TAT	CGA	GAC	ACC	2466
V	R	L	W	R	L	F	R	E	I	L	833
GTC	AGA	CTC	TGG	AGG	CTT	TTT	CGA	GAG	ATT	CTG	2499
D	G	L	A	Y	I	H	E	K	G	M	844
GAT	GGA	TTA	GCT	TAT	ATC	CAT	GAG	AAA	GGA	ATG	2532
I	H	R	D	L	K	P	V	N	I	F	855
ATT	CAC	CGG	GAT	TTG	AAG	CCT	GTC	AAC	ATT	TTT	2565
L	D	S	D	D	H	V	K	I	G	D	866
TTG	GAT	TCT	GAT	GAC	CAT	GTG	AAA	ATA	GGT	GAT	2598

FIG. 1F





7/53

F	G	L	A	T	D	H	L	A	F	S	877
TTT	GGT	TTG	GCG	ACA	GAC	CAT	CTA	GCC	TTT	TCT	2631
A	D	S	K	Q	D	D	Q	T	G	D	888
GCT	GAC	AGC	AAA	CAA	GAC	GAT	CAG	ACA	GGA	GAC	2664
L	I	K	S	D	P	S	G	H	L	T	899
TTG	ATT	AAG	TCA	GAC	CCT	TCA	GGT	CAC	TTA	ACT	2697
G	M	V	G	T	A	L	Y	V	S	P	910
GGG	ATG	GTT	GGC	ACT	GCT	CTC	TAT	GTA	AGC	CCA	2730
E	V	Q	G	S	T	K	S	A	Y	N	921
GAG	GTC	CAA	GGA	AGC	ACC	AAA	TCT	GCA	TAC	AAC	2763
Q	K	V	D	L	F	S	L	G	I	I	932
CAG	AAA	GTG	GAT	CTC	TTC	AGC	CTG	GGA	ATT	ATC	2796
F	F	E	M	S	Y	H	P	M	V	T	943
TTC	TTT	GAG	ATG	TCC	TAT	CAC	CCC	ATG	GTC	ACG	2829
A	S	E	R	I	F	V	L	N	Q	L	954
GCT	TCA	GAA	AGG	ATC	TTT	GTT	CTC	AAC	CAA	CTC	2862
R	D	P	T	S	P	K	F	P	E	D	965
AGA	GAT	CCC	ACT	TCG	CCT	AAG	TTT	CCA	GAA	GAC	2895
F	D	D	G	E	H	A	K	Q	K	S	976
TTT	GAC	GAT	GGA	GAG	CAT	GCA	AAG	CAG	AAA	TCA	2928
V	I	S	W	L	L	N	H	D	P	A	987
GTC	ATC	TCC	TGG	CTG	TTG	AAC	CAC	GAT	CCA	GCA	2961
K	R	P	T	A	T	E	L	L	K	S	998
AAA	CGG	CCC	ACA	GCC	ACA	GAA	CTG	CTC	AAG	AGT	2994
E	L	L	P	P	P	Q	M	E	E	S	1009
GAG	CTG	CTG	CCC	CCA	CCC	CAG	ATG	GAG	GAG	TCA	3027
E	L	H	E	V	L	H	H	T	L	T	1020
GAG	CTG	CAT	GAA	GTG	CTG	CAC	CAC	ACG	CTG	ACC	3060

FIG. 1G



8/53

APPROVED BY DRAFTSMAN	C.G. FIG.	SUBCLASS
	CLASS	

N	V	D	G	K	A	Y	R	T	M	M	1031
AAC	GTG	GAT	GGG	AAG	GCC	TAC	CGC	ACC	ATG	ATG	3093
A	Q	I	F	S	Q	R	I	S	P	A	1042
GCC	CAG	ATC	TTC	TCG	CAG	CGC	ATC	TCC	CCT	GCC	3126
I	D	Y	T	Y	D	S	D	I	L	K	1053
ATC	GAT	TAC	ACC	TAT	GAC	AGC	GAC	ATA	CTG	AAG	3159
G	N	F	S	I	R	T	A	K	M	Q	1064
GGC	AAC	TTC	TCA	ATC	CGT	ACA	GCC	AAG	ATG	CAG	3192
Q	H	V	C	E	T	I	I	R	I	F	1075
CAG	CAT	GTG	TGT	GAA	ACC	ATC	ATC	CGC	ATC	TTT	3225
K	R	H	G	A	V	Q	L	C	T	P	1086
AAA	AGA	CAT	GGA	GCT	GTT	CAG	TTG	TGT	ACT	CCA	3258
L	L	L	P	R	N	R	Q	I	Y	E	1097
CTA	CTG	CTT	CCC	CGA	AAC	AGA	CAA	ATA	TAT	GAG	3291
H	N	E	A	A	L	F	M	D	H	S	1108
CAC	AAC	GAA	GCT	GCC	CTA	TTC	ATG	GAC	CAC	AGC	3324
G	M	L	V	M	L	P	F	D	L	R	1119
GGG	ATG	CTG	GTG	ATG	CTT	CCT	TTT	GAC	CTG	CGG	3357
I	P	F	A	R	Y	V	A	R	N	N	1130
ATC	CCT	TTT	GCA	AGA	TAT	GTG	GCA	AGA	AAT	AAT	3390
I	L	N	L	K	R	Y	C	I	E	R	1141
ATA	TTG	AAT	TTA	AAA	CGA	TAC	TGC	ATA	GAA	CGT	3423
V	F	R	P	R	K	L	D	R	F	H	1152
GTG	TTC	AGG	CCG	CGC	AAG	TTA	GAT	CGA	TTT	CAT	3456
P	K	E	L	L	E	C	A	F	D	I	1163
CCC	AAA	GAA	CTT	CTG	GAG	TGT	GCA	TTT	GAT	ATT	3489

FIG. 1H

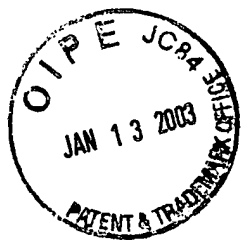


9/53

O.G. FIG.	SUBCLASS
	CLASS
APPROVED BY	DRAFTSMAN

V	T	S	T	T	N	S	F	L	P	T	1174
GTC	ACT	TCT	ACC	ACC	AAC	AGC	TTT	CTG	CCC	ACT	3522
A	E	I	I	Y	T	I	Y	E	I	I	1185
GCT	GAA	ATT	ATC	TAC	ACT	ATC	TAT	GAA	ATC	ATC	3555
Q	E	F	P	A	L	Q	E	R	N	Y	1196
CAA	GAG	TTT	CCA	GCA	CTT	CAG	GAA	AGA	AAT	TAC	3588
S	I	Y	L	N	H	T	M	L	L	K	1207
AGT	ATT	TAT	TTG	AAC	CAT	ACC	ATG	TTA	TTG	AAA	3621
A	I	L	L	H	C	G	I	P	E	D	1218
GCA	ATA	CTC	TTA	CAC	TGT	GGG	ATC	CCA	GAA	GAT	3654
K	L	S	Q	V	Y	I	I	L	Y	D	1229
AAA	CTC	AGT	CAA	GTC	TAC	ATT	ATT	CTG	TAT	GAT	3687
A	V	T	E	K	L	T	R	R	E	V	1240
GCT	GTG	ACA	GAG	AAG	CTG	ACG	AGG	AGA	GAA	GTG	3720
E	A	K	F	C	N	L	S	L	S	S	1251
GAA	GCT	AAA	TTT	TGT	AAT	CTG	TCT	TTG	TCT	TCT	3753
N	S	L	C	R	L	Y	K	F	I	E	1262
AAT	AGT	CTG	TGT	CGA	CTC	TAC	AAG	TTT	ATT	GAA	3786
Q	K	G	D	L	Q	D	L	M	P	T	1273
CAG	AAG	GGA	GAT	TTG	CAA	GAT	CTT	ATG	CCA	ACA	3819
I	N	S	L	I	K	Q	K	T	G	I	1284
ATA	AAT	TCA	TTA	ATA	AAA	CAG	AAA	ACA	GGT	ATT	3852
A	Q	L	V	K	Y	G	L	K	D	L	1295
GCA	CAG	TTG	GTG	AAG	TAT	GGC	TTA	AAA	GAC	CTA	3885
E	E	V	V	G	L	L	K	K	L	G	1306
GAG	GAG	GTT	GTT	GGA	CTG	TTG	AAG	AAA	CTC	GGC	3918

FIG. 1I



10/53

I	K	L	Q	V	L	I	N	L	G	L	1317
ATC	AAG	TTA	CAG	GTC	TTG	ATC	AAT	TTG	GGC	TTG	3951
V	Y	K	V	Q	Q	H	N	G	I	I	1328
GTT	TAC	AAG	GTG	CAG	CAG	CAC	AAT	GGA	ATC	ATC	3984
F	Q	F	V	A	F	I	K	R	R	Q	1339
TTC	CAG	TTT	GTG	GCT	TTC	ATC	AAA	CGA	AGG	CAA	4017
R	A	V	P	E	I	L	A	A	G	G	1350
AGG	GCT	GTA	CCT	GAA	ATC	CTC	GCA	GCT	GGA	GGC	4050
R	Y	D	L	L	I	P	Q	F	R	G	1361
AGA	TAT	GAC	CTG	CTG	ATT	CCC	CAG	TTT	AGA	GGG	4083
P	Q	A	L	G	P	V	P	T	A	I	1372
CCA	CAA	GCT	CTG	GGG	CCA	GTT	CCC	ACT	GCC	ATT	4116
G	V	S	I	A	I	D	K	I	S	A	1383
GGG	GTC	AGC	ATA	GCT	ATA	GAC	AAG	ATA	TCT	GCT	4149
A	V	L	N	M	E	E	S	V	T	I	1394
GCT	GTC	CTC	AAC	ATG	GAG	GAA	TCT	GTT	ACA	ATA	4182
S	S	C	D	L	L	V	V	S	V	G	1405
AGC	TCT	TGT	GAC	CTC	CTG	GTT	GTA	AGT	GTT	GGT	4215
Q	M	S	M	S	R	A	I	N	L	T	1416
CAG	ATG	TCT	ATG	TCC	AGG	GCC	ATC	AAC	CTA	ACC	4248
Q	K	L	W	T	A	G	I	T	A	E	1427
CAG	AAA	CTC	TGG	ACA	GCA	GGC	ATC	ACA	GCA	GAA	4281
I	M	Y	D	W	S	Q	S	Q	E	E	1438
ATC	ATG	TAC	GAC	TGG	TCA	CAG	TCC	CAA	GAG	GAA	4314
L	Q	E	Y	C	R	H	H	E	I	T	1449
TTA	CAA	GAG	TAC	TGC	AGA	CAT	CAT	GAA	ATC	ACC	4347
Y	V	A	L	V	S	D	K	E	G	S	1460
TAT	GTG	GCC	CTT	GTC	TCG	GAT	AAA	GAA	GGA	AGC	4380

FIG. 1J

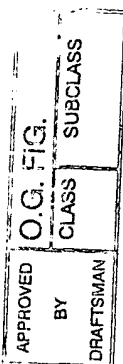
APPROVED	U.C. FIG.	SUBCLASS
BY	CLASS	
DRAFTSMAN		



11/53

H	V	K	V	K	S	F	E	K	E	R	1471
CAT	GTC	AAG	GTT	AAG	TCT	TTC	GAG	AAG	GAA	AGG	4413
Q	T	E	K	R	V	L	E	T	E	L	1482
CAG	ACA	GAG	AAG	CGT	GTG	CTG	GAG	ACT	GAA	CTT	4446
V	D	H	V	L	Q	K	L	R	T	K	1493
GTG	GAC	CAT	GTA	CTG	CAG	AAA	CTG	AGG	ACT	AAA	4479
V	T	D	E	R	N	G	R	E	A	S	1504
GTC	ACT	GAT	GAA	AGG	AAT	GGC	AGA	GAA	GCT	TCC	4512
D	N	L	A	V	Q	N	L	K	G	S	1515
GAT	AAT	CTT	GCA	GTG	CAA	AAT	CTG	AAG	GGG	TCA	4545
F	S	N	A	S	G	L	F	E	I	H	1526
TTT	TCT	AAT	GCT	TCA	GGT	TTG	TTT	GAA	ATC	CAT	4578
G	A	T	V	V	P	I	V	S	V	L	1537
GGA	GCA	ACA	GTG	GTT	CCC	ATT	GTG	AGT	GTG	CTA	4611
A	P	E	K	L	S	A	S	T	R	R	1548
GCC	CCG	GAG	AAG	CTG	TCA	GCC	AGC	ACT	AGG	AGG	4644
R	Y	E	T	Q	V	Q	T	R	L	Q	1559
CGC	TAT	GAA	ACT	CAG	GTA	CAA	ACT	CGA	CTT	CAG	4677
T	S	L	A	N	L	H	Q	K	S	S	1570
ACC	TCC	CTT	GCC	AAC	TTA	CAT	CAG	AAA	AGC	AGT	4710
E	I	E	I	L	A	V	D	L	P	K	1581
GAA	ATT	GAA	ATT	CTG	GCT	GTG	GAT	CTA	CCC	AAA	4743
E	T	I	L	Q	F	L	S	L	E	W	1592
GAA	ACA	ATA	TTA	CAG	TTT	TTA	TCA	TTA	GAG	TGG	4776
D	A	D	E	Q	A	F	N	T	T	V	1603
GAT	GCT	GAT	GAA	CAG	GCA	TTT	AAC	ACA	ACT	GTG	4809

FIG. 1K





12/53

K	Q	L	L	S	R	L	P	K	Q	R	1614
AAG	CAG	CTG	CTG	TCA	CGC	CTG	CCA	AAG	CAA	AGA	4842
Y	L	K	L	V	C	D	E	I	Y	N	1625
TAC	CTC	AAA	TTA	GTC	TGT	GAT	GAA	ATT	TAT	AAC	4875
I	K	V	E	K	K	V	S	V	L	F	1636
ATC	AAA	GTA	GAA	AAA	AAG	GTG	TCT	GTG	CTA	TTT	4908
L	Y	S	Y	R	D	D	Y	Y	R	I	1647
CTG	TAC	AGC	TAT	AGA	GAT	GAC	TAC	TAC	AGA	ATC	4941
L	F	*									1650
TTA	TTT	TAA									4950 ← SEQ.ID NO: 3

APPROVED BY DRAFTSMAN	O.G. FIG.	SUBCLASS
	CLASS	

CCCTAAAGAACTGTCGTTAACCTCATTCAAACAGACAGAGGCTTATACTG
GAATAATGGAATGTTGTACATTCATCATAATTTAAAATTAAATTCTAAGA
AGAGGCTGGGTGCAGTGGCTCACACCTTTAATCCCAGCACTTTGGGAAGC
CAAGGCAGGAAGACTGCTTGAAACCAGGAGTTTGAGACCAGCCTGAGCAA
CAAAGCAAGACCCCATCTCTATAAAAACTAAAAAAATTAGTTGGGCATGG
TGGCACATGCCTGTAGTCCCAGCTACTCCAGAGGCTGAGATGGATCATCT
GAGCCTCAGGAGGTTGAGGCTGCAGTGAGCTGTGACTGCGCCACTGCACT
CCAGTCTGGGACAACAGAGCAAGACCCTGTCTTAAAAAAAAAAGAAAAA
AAAATTTTTTTTTCTAAGAAGCTGTCCTACAAAGTTGAGCTTTGTTAGTTT
TTCATGTGTAATATATTATAAATTTATCTTTTGGGATATAATAAATGCTT
TCATATACCTGCA

FIG. 1L

6065914_eIF2kinase_man	-----
14790	MAGGRGAPGRGRDEPPESYPQRQDH
6066585_eIF2kinase_mouse	MAGGRGASGRGRAEPQESYSQRQDH
6065914_eIF2kinase_man	-----
14790	ELQALEAIYGADFQDLRPDACGPVK
6066585_eIF2kinase_mouse	ELQALEAIYGSDFQDLRPDARGRVR
6065914_eIF2kinase_man	-----
14790	EPPEINLVLYPQGLTGEEVYVKVDL
6066585_eIF2kinase_mouse	EPPEINLVLYPQGLAGEEVYVQVEL

FIG. 2A



13/53

6065914_eIF2kinase_man	-----
14790	RVKCPPTYPDVVPEIELKNAKGLSN
6066585_eIF2kinase_mouse	QVKCPPTYPDVVPEIELKNAKGLSN
6065914_eIF2kinase_man	-----
14790	ESVNLLKSRLEELAKKHCGEVMIFE
6066585_eIF2kinase_mouse	ESVNLLKSHLEELAKKQCGEVMIFE
6065914_eIF2kinase_man	-----
14790	LAYHVQSFLSEHNKPPPKSFHEEML
6066585_eIF2kinase_mouse	LAHHVQSFLSEHNKPPPKSFHEEML
6065914_eIF2kinase_man	-----
14790	ERRAQEEQQRLLEAKRKEEQEQREI
6066585_eIF2kinase_mouse	ERQAQEKQQRLLLEARRKEEQEQREI
6065914_eIF2kinase_man	-----
14790	LHEIQRRKEEIKKEKKRKEMAKQER
6066585_eIF2kinase_mouse	LHEIQRRKEEIKKEKKRKEMAKQER
6065914_eIF2kinase_man	-----
14790	LEIASLSNQDHTSKKDPGGHRTAAI
6066585_eIF2kinase_mouse	LEITSLTNQDYASKRDPAGHRAAAI
6065914_eIF2kinase_man	-----
14790	LHGGSPDFVGNGKHRANSSGRSRRE
6066585_eIF2kinase_mouse	LHGGSPDFVGNGKARTYSSGRSRRE
6065914_eIF2kinase_man	-----
14790	RQYSVCNSEDSPGSCEILYFNMGSP
6066585_eIF2kinase_mouse	RQYSVCSGEPSPGSCDILHFSVGSP
6065914_eIF2kinase_man	-----
14790	DQLMVHKGKKGKIGSDEQLGKLVYNAL
6066585_eIF2kinase_mouse	DQLMVHKGRCVGSDEQLGKVYNAL
6065914_eIF2kinase_man	-----
14790	ETATGGFVLLYEWVLQWQKKMGPFLL
6066585_eIF2kinase_mouse	ETATGSFVLLHEWVLQWQK-MGPCL

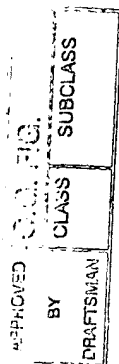
FIG. 2B



14/53

6065914_eIF2kinase_man	-----
14790	TSQEKEKIDKCKKQIQGTETEFNSL
6066585_eIF2kinase_mouse	TSQEKEKIDKCKRQIQGAETEFSSL
6065914_eIF2kinase_man	-----
14790	VKLSHPNVVRYLAMNLKEQDDSIIV
6066585_eIF2kinase_mouse	VKLSHPNIVRYFAMNSREEEDSIVI
6065914_eIF2kinase_man	-----
14790	DILVEHISGVSLAAHLSHSGPIPVH
6066585_eIF2kinase_mouse	DILAEHVSGISLATHLSHSGPVPAP
6065914_eIF2kinase_man	-----
14790	QLRRYTAQLLSGLDYLHSNSVVHKV
6066585_eIF2kinase_mouse	QLRKYTAQLLAGLDYLHSNSVVHKV
6065914_eIF2kinase_man	-----
14790	LSASNVLVDAEGTVKITDYSISKRL
6066585_eIF2kinase_mouse	LSASSVLVDAEGTVKITDYSISKRL
6065914_eIF2kinase_man	-----
14790	ADICKEDVFEQTRVRFSDNALPYKT
6066585_eIF2kinase_mouse	ADICKEDVFEQARVRFSDSALPYKT
6065914_eIF2kinase_man	-----
14790	GKKGDVWRLGLLLLSLSQGQECGEY
6066585_eIF2kinase_mouse	GKKGDVWRLGLLLLSLSQGQECGEY
6065914_eIF2kinase_man	-----
14790	PVTIPSDLPADFQDFLKKCVCLDDK
6066585_eIF2kinase_mouse	PVTIPSDLPADFQDFLKKCVCLDDK
6065914_eIF2kinase_man	-----
14790	ERWSPQQLLKHSFINPQPKMPLVEQ
6066585_eIF2kinase_mouse	ERWSPQQLLKHSFINPQPKLPLVEQ
6065914_eIF2kinase_man	-----
14790	SPEDSGGQDYVETVIPSNRLPSAAF

FIG. 2C





15/53

APPROVED BY DRAFTSMAN	O.G. FIG.	SUBCLASS
	CLASS	

6066585_eIF2kinase_mouse SPEDSGGQDYIETVIPS NQLPSAAF

6065914_eIF2kinase_man -----
14790 FSETQRQFSRYFIEFEELQLLGKGA
6066585_eIF2kinase_mouse FSETQKQFSRYFIEFEELQLLGKGA

6065914_eIF2kinase_man -----
14790 FGAVIKVQNKLDGCCYAVKRIPINP
6066585_eIF2kinase_mouse FGAVIKVQNKLDGCCYAVKRIPINP

6065914_eIF2kinase_man -----
14790 ASRQFRRIKGEVTLLSRLHHENIVR
6066585_eIF2kinase_mouse ASRHFRIKGEVTLLSRLHHENIVR

6065914_eIF2kinase_man -----
14790 YYNWIERHERPAGPGT PPPDSGPL
6066585_eIF2kinase_mouse YYNWIERHERPAVPGT PPPDCTPQ

6065914_eIF2kinase_man -----
14790 AKDDRAARGQPASDTDGLDSVEAAA
6066585_eIF2kinase_mouse AQDSPATCGKTS GDTEELGSVEAAA

6065914_eIF2kinase_man -----
14790 PPILSSSVEWSTSGERSASARFPA
6066585_eIF2kinase_mouse PPILSSSVEWSTSAERSTSTRFPV

6065914_eIF2kinase_man -----
14790 TGPSSDDEDDDEDEHGGVFSQSFL
6066585_eIF2kinase_mouse TGQDSSSDEED-EDERDGVFSQSFL

6065914_eIF2kinase_man -----
14790 PASDSESDIIFDNEDENSKSQNQDE
6066585_eIF2kinase_mouse PASDSDSDIIFDNEDENSKSQNQDE

6065914_eIF2kinase_man -----
14790 DCNEKNGCHESEPSVTTEAVHYLYI

FIG. 2D



16/53

6066585_eIF2kinase_mouse DCNQKDGSHIEIEPSVTAEAVHYLYI

6065914_eIF2kinase_man -----
14790 QMEYCEKSTLRDTIDQGLYRDTVRL
6066585_eIF2kinase_mouse QMEYCEKSTLRDTIDQGLFRDTSRL

6065914_eIF2kinase_man -----
14790 WRLFREILDGLAYIHEKGMHRDLK
6066585_eIF2kinase_mouse WRLFREILDGLAYIHEKGMHRDLK

6065914_eIF2kinase_man -----
14790 PVNIFLDSDDHVKIGDFGLATDHLA
6066585_eIF2kinase_mouse PVNIFLDSDDHVKIGDFGLATDHLA

6065914_eIF2kinase_man -----
14790 FSADSKQDDQTGD-LIKSDPSGHLT
6066585_eIF2kinase_mouse FTAEGKQDDQAGDGVIKSDPSGHLT

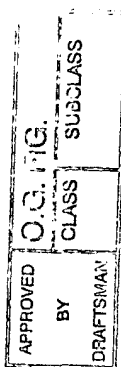
6065914_eIF2kinase_man -----
14790 GMVGTALYVSPEVQGSTKSAYNQKV
6066585_eIF2kinase_mouse GMVGTALYVSPEVQGSTKSAYNQKV

6065914_eIF2kinase_man -----
14790 DLFSLGIIFFEMSYHPMVTASERIF
6066585_eIF2kinase_mouse DLFSLGIIFFEMSYHPMVTASERIF

6065914_eIF2kinase_man -----
14790 VLNQLRDPTSPKFPEDFDDGEHAKQ
6066585_eIF2kinase_mouse VLNQLRDPTSPKFPDDFDDGEHTKQ

6065914_eIF2kinase_man -----
14790 KSVISWLLNHDPAPKRPTATELLKSE
6066585_eIF2kinase_mouse KSVISWLLNHDPAPKRPTAMELLKSE

FIG. 2E





17/53

6065914_eIF2kinase_man -----
14790 LLPPPQMEESSELHEVLHHTLTNVDG
6066585_eIF2kinase_mouse LLPPPQMEESSELHEVLHHTLANIDG

6065914_eIF2kinase_man -----
14790 KAYRTMMAQIFSQRISPAIDYTYDS
6066585_eIF2kinase_mouse KAYRTMMSQIFCQHISPAIDYTYDS

6065914_eIF2kinase_man -----
14790 DILKGNFSIRTAKMQQHVCETIIRI
6066585_eIF2kinase_mouse DILKGNFLIRTAKIQQLVCETIVRV

6065914_eIF2kinase_man -----
14790 FKRHGAVQLCTPLLLPRNRQIYEHN
6066585_eIF2kinase_mouse FKRHGAVQLCTPLLLPRNRQIYEHN

6065914_eIF2kinase_man --ALFMDHSGMLVMLPFDLRIPFAR
14790 EAALFMDHSGMLVMLPFDLRIPFAR
6066585_eIF2kinase_mouse EAALFMDHSGMLVMLPFDLRVPFAR
*****:****

6065914_eIF2kinase_man YVARNNILILKRYCIERVFRPRKLD
14790 YVARNNILNLKRYCIERVFRPRKLD
6066585_eIF2kinase_mouse YVARNNILNLKRYCIERVFRPRKLD

6065914_eIF2kinase_man RFHPKELLECAFDIVTSTTNSFLPT
14790 RFHPKELLECAFDIVTSTTNSFLPT
6066585_eIF2kinase_mouse RFHPKELLECAFDIVTSTTNSLPT

6065914_eIF2kinase_man AEIIYTIYEIIQEFPALQERNYSIY
14790 AEIIYTIYEIIQEFPALQERNYSIY
6066585_eIF2kinase_mouse AETIYTIYEIIQEFPALQERNYSIY
** *****

FIG. 2F



18/53

6065914_eIF2kinase_man LNHTMLLKAILLHCGIPEDKLSQVY
14790 LNHTMLLKAILLHCGIPEDKLSQVY
6066585_eIF2kinase_mouse LNHTMLLKAILLHCGIPEDKLSQVY

6065914_eIF2kinase_man IILYDAVTEKLTRREVEAKFCNLSL
14790 IILYDAVTEKLTRREVEAKFCNLSL
6066585_eIF2kinase_mouse VILYDAVTEKLTRREVEAKFCNLSL
:*****

6065914_eIF2kinase_man SSNSLCRLYKFIEQKGDLDLMPTI
14790 SSNSLCRLYKFIEQKGDLDLMPTI
6066585_eIF2kinase_mouse SSNSLCRLYKFIEQKGDLDLTPTI
***** ***

6065914_eIF2kinase_man NSLIKQKTGIAQLVKYGLKDLEEVV
14790 NSLIKQKTGIAQLVKYGLKDLEEVV
6066585_eIF2kinase_mouse NSLIKQKTGVAQLVKYSLKDLEDVV
:*****:*****.*****:**

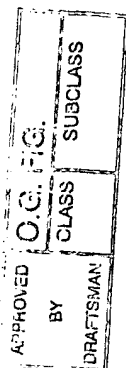
6065914_eIF2kinase_man GLLKKLGIKLQVLINLGLVYKVQQH
14790 GLLKKLGIKLQVLINLGLVYKVQQH
6066585_eIF2kinase_mouse GLLKKLGVKLQVSINLGLVYKVQQH
*****:***** ***** ***

6065914_eIF2kinase_man NGIIFQFVAFIKRRQRAVPEILAAG
14790 NGIIFQFVAFIKRRQRAVPEILAAG
6066585_eIF2kinase_mouse TGIIFQFLAFSKRRQRVVPEILAAG
.*****:** *****.*****

6065914_eIF2kinase_man GRYDLLIPQFRGPQALGPVPTAIGV
14790 GRYDLLIPQFRGPQALGPVPTAIGV
6066585_eIF2kinase_mouse GRYDLLIPKFRGPQTVGPVPTAVGV
*****:*****.:*****:**

6065914_eIF2kinase_man SIAIDKISAAVLNMEESVTISSCDL
14790 SIAIDKISAAVLNMEESVTISSCDL
6066585_eIF2kinase_mouse SIAIDKIFAVVLNMEEPVTVSSCDL
***** *.*****.***:*****

FIG. 2G





19/53

6065914_eIF2kinase_man LVVSVGQMSMSRAINLTQKLWTAGI
14790 LVVSVGQMSMSRAINLTQKLWTAGI
6066585_eIF2kinase_mouse LVVSVGQMSMSRAINLTQKLWTAGI

6065914_eIF2kinase_man TAEIMYDWSQSQEELQEYCRHHEIT
14790 TAEIMYDWSQSQEELQEYCRHHEIT
6066585_eIF2kinase_mouse TAEIMYDWSQSQEELQEYCRHHEIT

6065914_eIF2kinase_man YVALVSDKEGSHVKVKSFEKERQTE
14790 YVALVSDKEGSHVKVKSFEKERQTE
6066585_eIF2kinase_mouse YVALVSDKEGSHVKVKSFEKERQTE

6065914_eIF2kinase_man KRVLETELVLDHVLQKLRTKVTDERN
14790 KRVLETELVLDHVLQKLRTKVTDERN
6066585_eIF2kinase_mouse KRVLESDLVDHVMQKLRTKVGDERN
*****:*****:***** *****

6065914_eIF2kinase_man GREASDNLAVQNLKGSFSNASGLFE
14790 GREASDNLAVQNLKGSFSNASGLFE
6066585_eIF2kinase_mouse FRDASDNLAVQTLKGSFSNASGLFE
*:*****.*****

6065914_eIF2kinase_man IHGATVVPIVSVLAPEKLSASTRRR
14790 IHGATVVPIVSVLAPEKLSASTRRR
6066585_eIF2kinase_mouse IHGTTVVPNVIVLAPEKLSASTRRR
:** * *****

6065914_eIF2kinase_man YETQVQTRLQTS LANLHQSSEIEI
14790 YETQVQTRLQTS LANLHQSSEIEI
6066585_eIF2kinase_mouse HEIQVQTRLQTTLANLHQSSEIEI
:* *****:*****

FIG. 2H



20/53

6065914_eIF2kinase_man LAVDLPKETILQFLSLEWDADEQAF
14790 LAVDLPKETILQFLSLEWDADEQAF
6066585_eIF2kinase_mouse LAVDLPKETILQFLSLEWDADEQAF
:** * *****

6065914_eIF2kinase_man NTTVKQLLSRLPKQRYLKLVCDEIY
14790 NTTVKQLLSRLPKQRYLKLVCDEIY
6066585_eIF2kinase_mouse NTTVKQLLSRLPKQRYLKLVCDEIY

6065914_eIF2kinase_man NIKVEKKVSVLFLYSYRDDYYRILF
14790 NIKVEKKVSVLFLYSYRDDYYRILF
↑SEQ. ID NO: 4
6066585_eIF2kinase_mouse NIKVEKKVSVLFLYSYRDDYYRILF
↑SEQ. ID NO: 5

FIG. 2I

APPROVED	BY	DRAFTSMAN
C.G. FIG	CLASS	SUBCLASS



21/53

APPROVED	O.G. FIG.	
BY	CLASS	SUBCLASS
DRAFTSMAN		

Ser/thr Kinase 14790 Expression on HBV + Liver

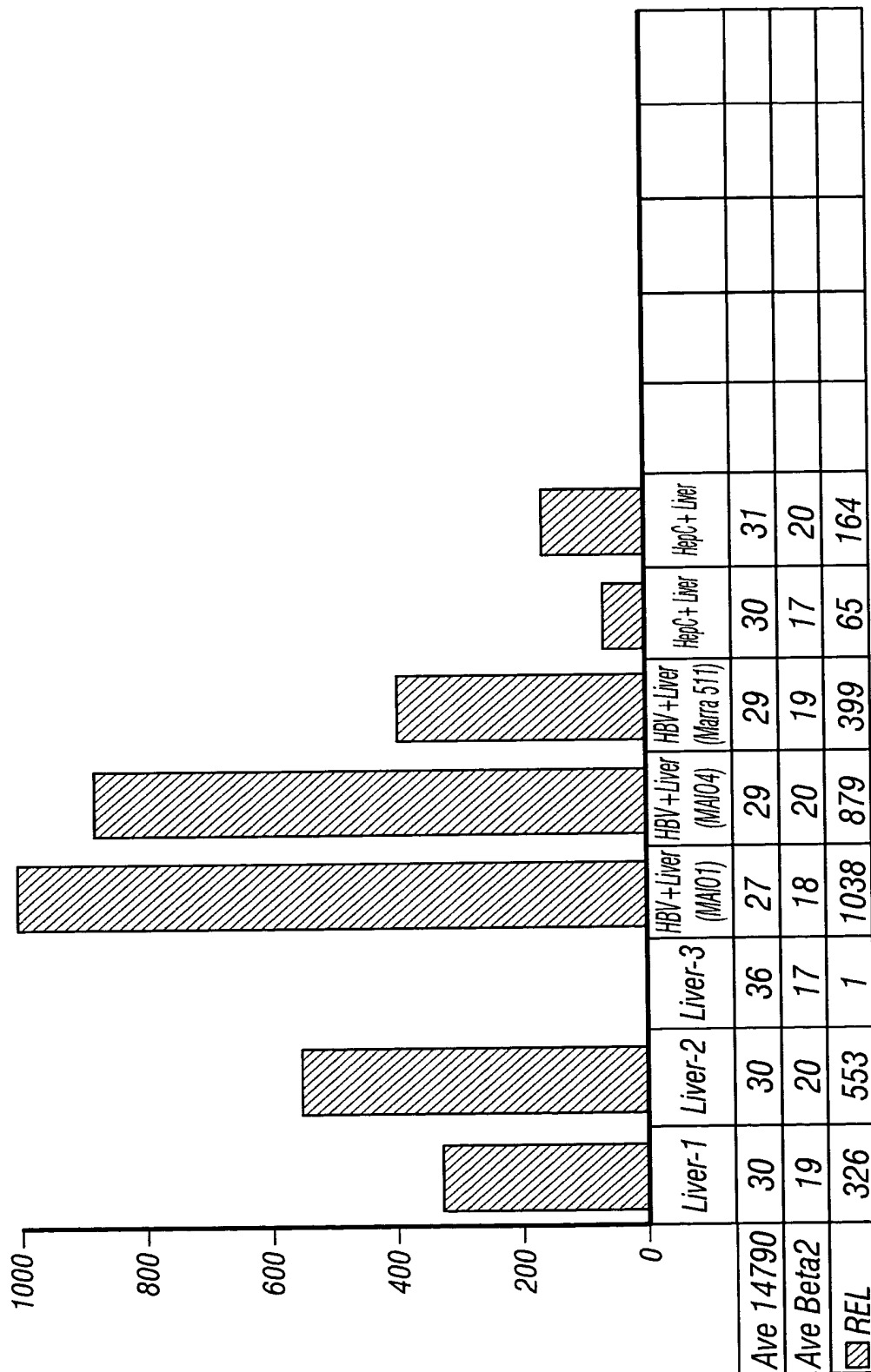


FIG. 3

Relative Expression (Liver PIT 260 used as reference sample)

APPROVED	O.G. FIG.	
BY	CLASS	SUBCLASS
DRAFTSMAN		



22/53

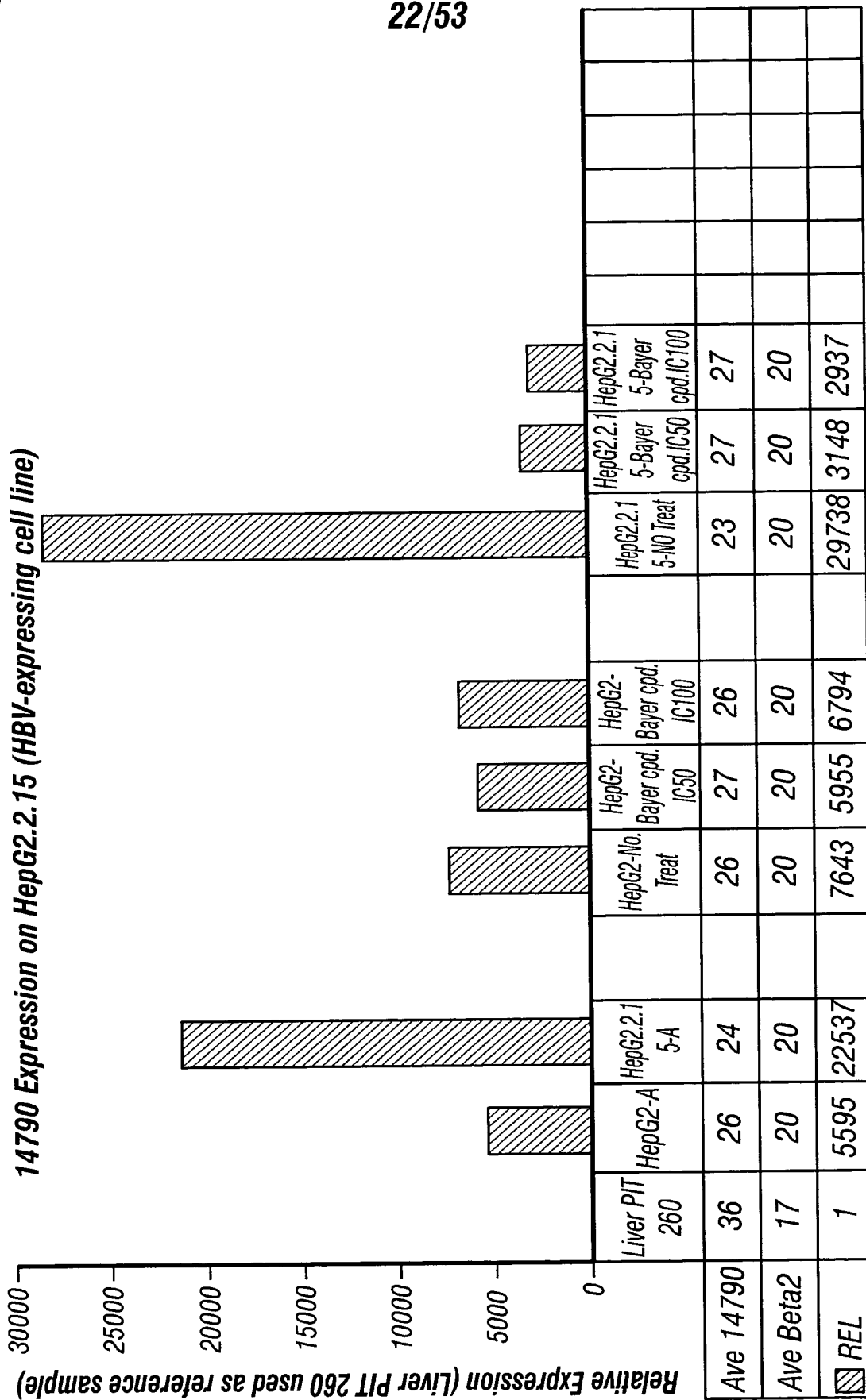


FIG. 4

APPROVED	O.G. FIG.	
BY	CLASS	SUBCLASS
DRAFTSMAN		



23/53

**Gene 14790 Expression in Normal Human Tissue
Relative Expression (Thyroid as Reference)**

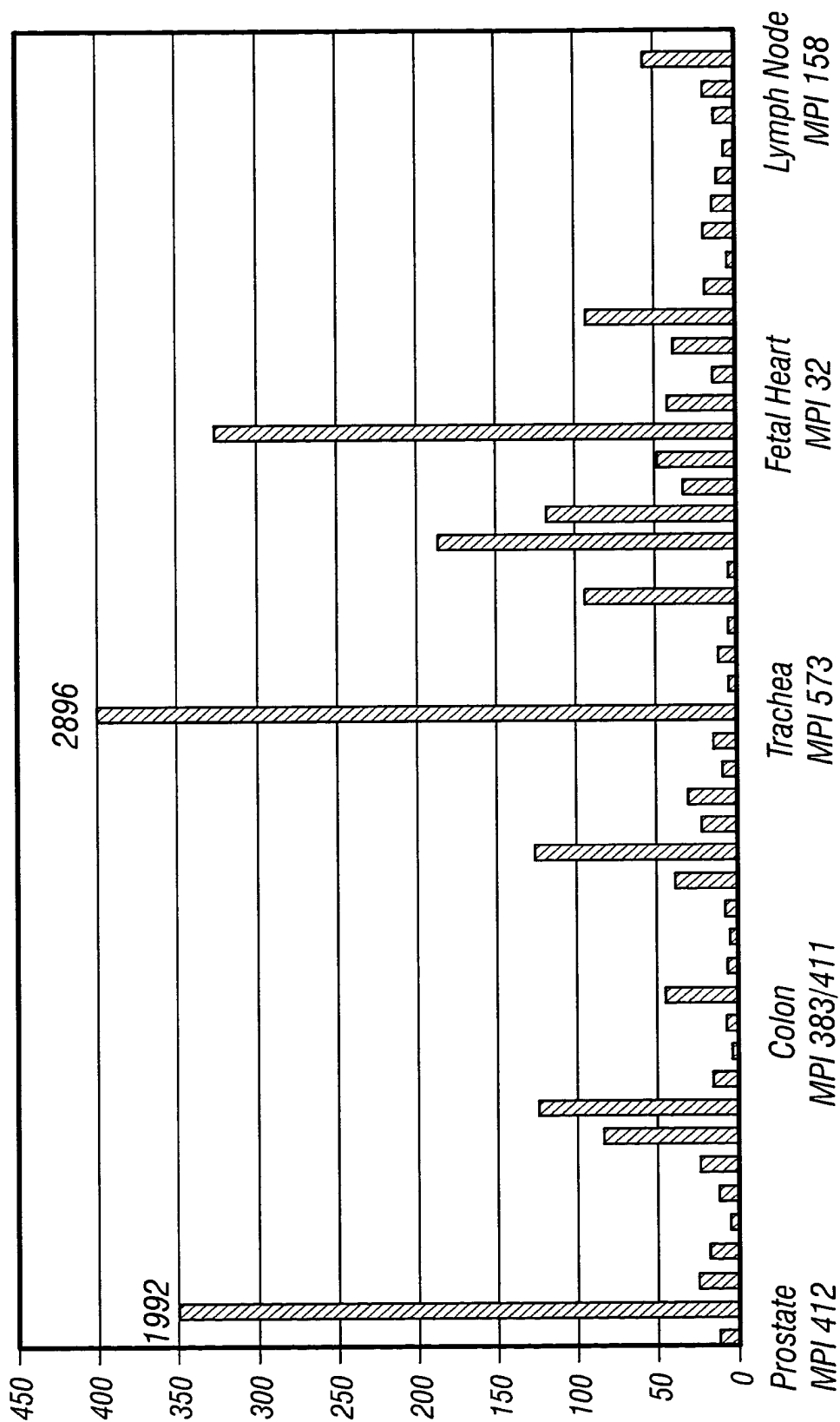


FIG. 5



24/53

APPROVED	BY	CLASS	SUBCLASS
	DRAFTSMAN		

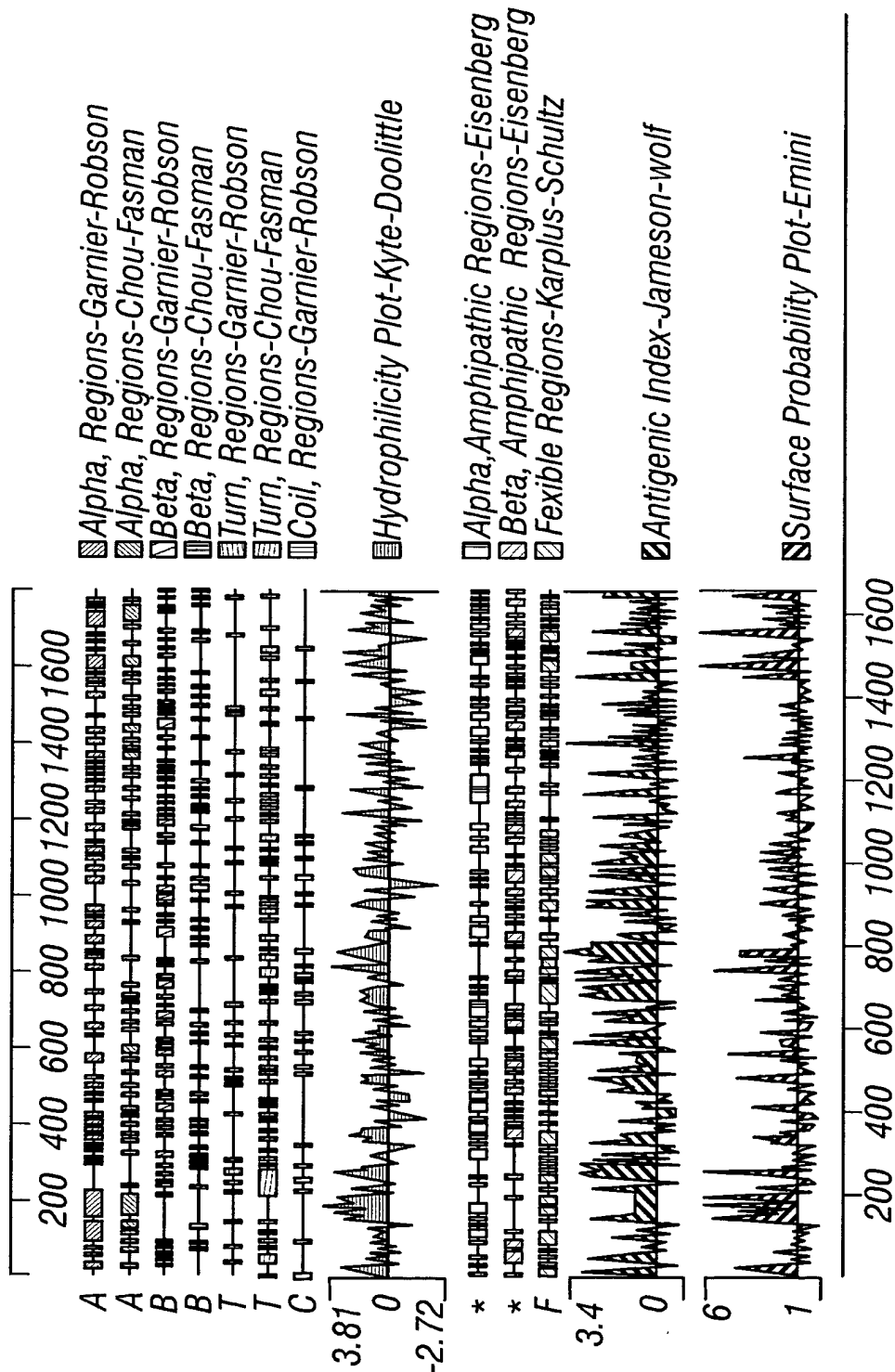
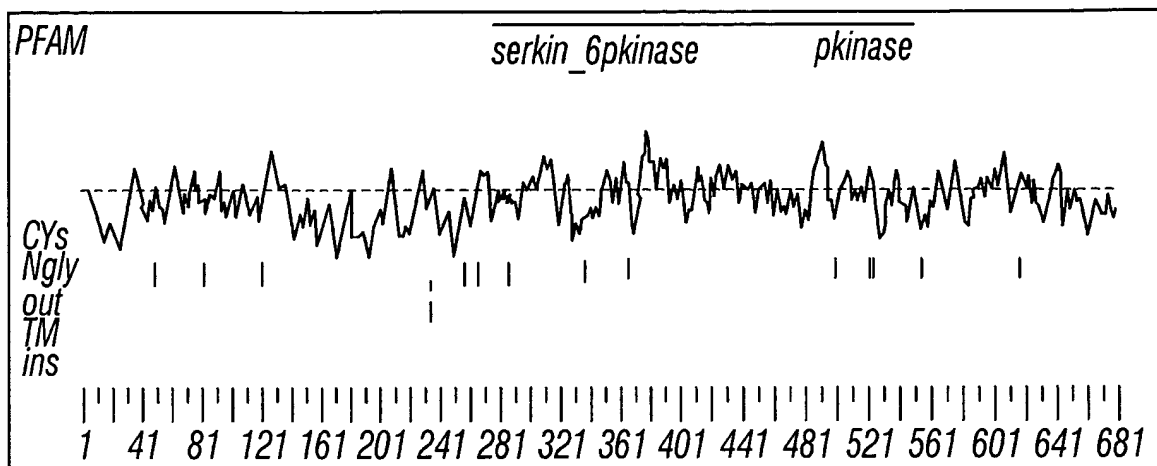


FIG. 6



25/53

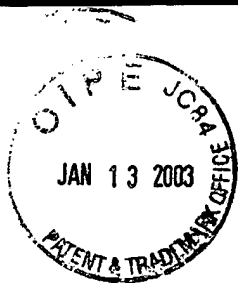
Analysis of 14790 (1649 aa)



>14790

MAGGRGAPGRGRDEPPESYPQRQDHELQALEAIYGADFQDLRPDACGPV
KEPPEINLVLYPQGLTGEEVYVKVDLRVKCPPTYPDVVPEIELKNAKGL
SNESVNLLKSRLEELAKKHCGEVMIFELAYHVQSFLSEHNKPPPKSFHE
EMLERRAQEEQQRLLLEAKRKEEQEQREILHEIQRRKEEIKEEKKRKEMA
KQERLEIASLSNQDHTSKKDPGGHRTAAILHGGSPDFVGNGKHRANSSG
RSRRERQYSVCNSEDSPGSCEILYFNMGSPDQLMVHKGKCIQSDEQLGK
LVYNALETATGGFVLLYEWVLQWQKKMGPFLLTSQEKEKIDKCKKQIQGT
ETEFNSLVKLSHPNVVRYLAMNLKEQDDSI VVDILVEHISGVSLAAHLS
HSGPIPVHQLRRYTAQLLSGLDYLHSNSVVKVLSASNVLVDAEGTVKI
TDYSISKRLADICKEDVFEQTRVRFSDNALPYKTGKKGDVWRLGLLLLS
LSQGQECGEY

FIG. 7A



26/53

APPROVED BY DRAFTSMAN	O.G. FIG.	SUBCLASS
	CLASS	

SPEDSGGQDYVETVIPS NRLPSAAFFSETQRQFSRYFIEFEELQLLGKGA
FGAVIKVQNKLDGCCYAVKRIPINPASRQFRRIKGEVTLLSRLHHENIVR
YYNAWIERHERPAGPGTPPPDSGPLAKDDRAARGQPASDTDGLDSVEAAA
PPILSSSVIEWSTSGERSASARFPATGPGSSDDEDDDEDEHGGVFSQSFL
PASDSESDIIFDNEDENSKSQNQDEDCNEKNGCHESEPSVTTEAVHYLYI
QMEYCEKSTLRDTIDQGLYRDTVRLWRLFREILDGLAYIHEKGMHRDLK
PVNIFLDSDDHVKIGDFGLATDHLAFSADSKQDDQTGDLIKSDPSGHLTG
MVG TALYVSPEVQGSTKSAYNQKVDLFSLGIIFFEMSYHPMVTASERIFV
LNQLRDPTSPKFPEDFDDGEHAKQKSVISWLLNHDPKRPTATELLKSEL
LPPPQMEESELHEVLHHTLTNVDGKAYRTMMAQIFSQRISPAIDYTYDSD
ILKGNFSIRTAKMQQHVCETIIRIFKRHGAVQLCTPLLLPRNRQIYEHNE
AALFMDHSGMLVMLPFDLRI PFARYVARNNILNLKRYCIE RVFRPRKLDR
FHPKELLECAFDIVTSTTNSFLPTAEIIYTIYEIIQEFPALQERNYSIYL
NHTMLLKAILLHCGIPEDKLSQVYIILYDAVTEKLTRREVEAKFCNLSLS
SNSLCRLYKFIEQKGDLDLMP TINS LIKQKTGIAQLVKYGLKDLEEVVG
LLKKLG IKLQVLINLGLVYKVQQHNGIIFQFVAFIKRRQRAVPEILAAGG
RYDLLIPQFRGPQALGPVPTAIGVSI AIDKISA AVLNMEE SVTISSCDLL
VVS VGQMSMSRAINLTQKLWTAGITAEIMYDWSQSQEELQEYCRHHEITY
VALVSDKEGSHVKVKSFEKERQTEKRVLET ELVDHVLQKLRTKVTDERNG
REASDNLAVQNLKGSFSNASGLFEIHGATVVPIVSVLAPEKLSASTRRRY
ETQVQTRLQTS LANLHQKSSEIEILAVDLPKETILQFLSLEWDADEQAFN
TTVKQLLSRLPKQRYLKLVCDEIYNIKVEKKVSVLFLYSYRDDYYRILF

PSORT Prediction of Protein Localization

MITDISC: discrimination of mitochondrial
targeting seq

R content:	3	Hyd Moment (75):	7.37
Hyd Moment (95):	6.02	G content:	5
D/E content:	2	S/7 content:	0

Gavel: prediction of cleavage sites for
mitochondrial preseq

R-2 motif at 20 GRG:RD

MUCDISC: discrimination of nuclear localization
signals

pat4: KK RK (5) at 190

FIG. 7B



27/53

pat4: RPRK (4) at 1144
pat7: none
bipartite: RRAQEEQQRLLEAKRKE at 152
bipartite: KRIPINPASRQFRRIKG at 619
content of basic residues: 12:11
NLS Score: 1.08

ER Membrane Retention Signals:

XXRR-like motif in the N-terminus: AGGR

None

Final Results (k = 9/23):

56.5 %: nuclear
30.4 %: cytoplasmic
4.3 %: vacuolar
4.3 %: mitochondrial
4.3 %: vesicles of secretory system

prediction for 14790 is nuc (k=23)

Start	End	Feature	Seq
144	204	coiled coil	FHEEMLERRA...AKQERLEIAS

Signal Peptide Predictions for 14790

Method	Predict	Score	Mat@
SignalP (eukaryote)	NO		

Note: amino-terminal 70aa used for signal peptide prediction

Transmembrane Segments Predicted by MEMSAT

Start	End	Orient	Score
1522	1538	ins-->out	0.8

FIG. 7C



28/53

>14790

MAGGRGAPGRGRDEPPESYPQRQDHELQALEAIYGADFQDLRPDACGPVK
EPPEINLVLYPQGLTGEEVYVKVDLRVKCPPTYPDVVPEIELKNAKGLSN
ESVNLLKSRLEELAKKHCGEVMIFELAYHVQSFLSEHNKPPPKSFHEEML
ERRAQEEQQRLLEAKRKEEQEQREILHEIQRRKEEIKKEKKRKEMAKQER
LEIASLSNQDHTSKKDPGGHRTAAILHGGSPDFVGNKGHRANSSSGRSRRE
RQYSVCNSEDSPGSC EILYFNMGSPDQLMVHKGK CIGSDEQLGKLVYNAL
ETATGGFVLLYEWVLQWQKMGPF LTSQEKEKIDKCKKQIQGTETEFNSL
VKLSHPNVVRYLAMNLKEQDDSI VVDILVEHISGVSLAAHLSHSGPIPVH
QLRRYTAQLLSGLDYLSNSVHVHKLVSASNVLVDAEGTVKITDYSISKRL
ADICKEDVFEQTRVRFSDNALPYKTGKKGDVWRLGLLLLSLSQGQECGEY
PVTIPSDLPADFQDFLKKCVCLDDKERWSPQQLLKHSFINPQPKMPLVEQ
SPEDSGGQDYVETVIPS NRLPSAAFFSETQRQFSRYFIEFEELQLLGKGA
FGAVIKVQNKLDGCCYAVKRI PINPASRQFRRIKGEVTLLSRLHHENIVR
YYNAWIERHERPAGPGTPPPDSGPLAKDDRAARGQPASDTDGLDSVEAAA
PPILSSSVIEWSTSGERSASARFPATGPGSSDDEDDDEDEHGGVFSQSFL
PASDSESDIIFDNEDENSKSQNQDEDCNEKNGCHESEPSVTTEAVHYLYI
QMEYCEKSTLRDTIDQGLYRDTVRLWRLFREILDGLAYIHEKGM IHRDLK
PVNIFLDSDDHVKIGDFGLATDHLAFSADSKQDDQTGD LIKSDPSGHLTG
MVG TALYVSPEVQGSTKSAYNQKVDLFSLG IIFFEMSYHPMVTASERIFV
LNQLRDP TSPKFPEDFDDGEHAKQKSVISWLLNHDP AKRPTATELLKSEL
LPPPQMEESELHEVLHHTLTNV DGKAYRTMMAQIFSQRISP AIDYTYDSD
ILKGNFSIRTAKMQQHVCETIIRIFKRHGAVQLCTPLLLPRNRQIYEHNE
AALFMDHSGMLVMLPFDLRIPFARYVARNNILNLKRYCIERVFRPRK LDR
FHPKELLECAFDIVTSTTNSFLPTAEI IYTIYEIIQEFPALQERNYSIYL
NHTMLLKAILLHCGIPEDKLSQVYIILYDAVTEKLTRREVEAKFCNLSLS
SNSLCRLYKFIEQKGD LQDLMPTINSLIKQKTGIAQLVKYGLKDLEEVVG
LLKKLG IKLQVLINLGLVYKVQQHNGIIFQFVAFIKRRQRAVPEILAAGG
RYDLLIPQFRGPQALGPVPTAIGVSI AIDKISAAVLNMEESVTISSCDLL
VVSVGQMSMSRAINLTQKLW TAGITAEIMYDWSQSQEELQEYCRHHEITY
VALVSDKEGSHVKVKSFEKERQTEKRVLET ELVDHVLQKLRTKVTDERNG
READSNLAVQNLKGSFSNASGLFEIHGATVVP IVSVLAPEKLSASTRRRY
ETQVQTRLQ TSLANLHQKSSEIEILAVDLPKETILQFLSLEWDADEQAFN
TTVKQLLSRLPKQRYLKLVCDEIYNIKVEKKVSVLFLYSYRDDYRILF

FIG. 7D



29/53

Prosites Pattern Matches for 14790

Prosites version: Release 12.2 of February 1995

>PS00021|PDOC00001|ASN_GLYCOSYLATION N-glycosylation site.

Query:	100	NESV	103
Query:	242	NSSG	245

Query:	1055	NFSI	1058
Query:	1195	NYSI	1198
Query:	1201	NHTM	1204
Query:	1246	NLSL	1249
Query:	1414	NLTQ	1417
Query:	1518	NASG	1521
Query:	1600	NTTV	1603

>PS00004|PDOC00004|CAMP_PHOSPHO_SITE cAMP- and cGMP-dependent protein kinase phosphorylation site.

Query:	403	RRYT	406
Query:	988	KRPT	991
Query:	1630	KKVS	1633

>PS00005|PDOC00005|PKC_PHOSPHO_SITE protein kinase C phosphorylation site.

Query:	212	TSK	214
Query:	244	SGR	246
Query:	247	SRR	249
Query:	438	TVK	440
Query:	447	SKR	449
Query:	475	TGK	477
Query:	567	SNR	569
Query:	579	TQR	581
Query:	720	SAR	722
Query:	809	TLR	811
Query:	822	TVR	824
Query:	915	STK	917
Query:	945	SER	947
Query:	959	SPK	961

FIG. 7E

O.G. FIG.	SUBCLASS
	CLASS
APPROVED BY DRAFTSMAN	



30/53

APPROVED BY DRAFTSMAN	O.G. FIG.	SUBCLASS
	CLASS	

Query: 1036 SQR 1038
Query: 1057 SIR 1059
Query: 1060 TAK 1062
Query: 1232 TEK 1234
Query: 1236 TRR 1238
Query: 1416 TQK 1418
Query: 1455 SDK 1457
Query: 1473 TEK 1475
Query: 1545 STR 1547
Query: 1602 TVK 1604
Query: 1639 SYR 1641

>PS00006|PDOC00006|CK2_PHOSPHO_SITE Casein kinase
II phosphorylation site.

Query: 65 TGEE 68
Query: 82 TYPD 85
Query: 108 SRLE 111
Query: 144 SFHE 147
Query: 207 SNQD 210
Query: 213 SKKD 216
Query: 247 SRRE 250
Query: 326 TSQE 329
Query: 343 TETE 346
Query: 411 SGLD 414
Query: 551 SPED 554
Query: 688 SDTD 691
Query: 707 SSVE 710
Query: 713 TSGE 716
Query: 730 SSDD 733
Query: 753 SDSE 756
Query: 809 TLRD 812
Query: 880 SKQD 883
Query: 943 TASE 946
Query: 991 TATE 994
Query: 1020 TNVD 1023
Query: 1180 TIYE 1183
Query: 1236 TRRE 1239

FIG. 7F



31/53

Query: 1395 SSCD 1398
Query: 1435 SQEE 1438
Query: 1455 SDKE 1458
Query: 1570 SEIE 1573
Query: 1639 SYRD 1642

APPROVED	BY	CLASS	SUCCESS
	DRAFTSMAN		

>PS00007|PDOC00007|TYR_PHOSPHO_SITE Tyrosine kinase phosphorylation site.

Query: 246 RSRRERQY 253
Query: 811 RDTIDQGLY 819
Query: 830 REILDGLAY 838
Query: 1444 RHHEITY 1450
Query: 1617 KLVCDEIY 1624

>PS00008|PDOC00008|MYRISTYL N-myristoylation site.

Query: 218 GGHRTA 223
Query: 384 GVSLAA 389
Query: 494 CQECGE 499
Query: 599 GAFGAV 604
Query: 613 GCCYAV 618
Query: 684 GQPASD 689

Query: 742 GGVFSQ 747
Query: 782 GCHESE 787
Query: 900 GMVGTA 905
Query: 914 GSTKSA 919
Query: 1373 GVSIAI 1378
Query: 1514 GSFSNA 1519

>PS00009|PDOC00009|AMIDATION Amidation site.

Query: 475 TGKK 478

>PS00107|PDOC00100|PROTEIN_KINASE_ATP Protein kinases ATP-binding region signature.



32/53

Query: 596 LGKGAFGAV 604

>PS00108|PDOC00100|PROTEIN_KINASE_ST
Serine/Threonine protein kinases active-site
signature.

Query: 844 MIHRDLKPVNIFL 856

>PS00116|PDOC00107|DNA_POLYMERASE_B DNA polymerase
family B signature.

Query: 687 ASDTDGLDS 695

Protein Family / Domain Matches, HMMer version 2

Searching for complete domains in PFAM
Hmmpfam search a single seq against HMM database
HMMER: 2.1.1 (Dec 1998)
Copyright (C) 1992-1998 Washington University
School of Medicine
HMMER is freely distributed under the GNU General
Public License (GPL).

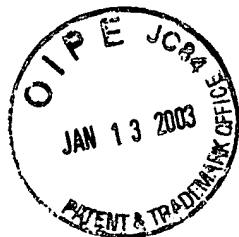
HMM file: /prod/ddm/seqanal/PFAM/pfam4.4/Pfam
Sequence file: /prod/ddm/wspace/orfanal/oa-
script.16895.seq

Query: 14790

Scores for sequence family classification (score
includes all domains):

Model	Description
-----	-----
<u>pkinase</u>	Eukaryotic protein kinase domain
<u>Ribosomal L23</u>	Ribosomal protein L23
<u>mRNA_cap_enzyme</u>	mRNA capping enzyme

FIG. 7H



33/53

	Score	E-value	N
<u>pkinese</u>	282.0	7.8e-81	4
<u>Ribosomal L23</u>	5.0	3.9	1
<u>mRNA_cap_enzyme</u>	-181.3	9.6	1

Parsed for domains:

Model	Domain	seq-f	seq-t	hmm-f	hmm-t
pkinese	1/4	332	443	.. 30	134..
pkinese	2/4	501	539	.. 237	278.]
pkinese	3/4	590	662	.. 1	66 [.
pkinese	4/4	797	1001	.. 65	278.]
Ribosomal_L23	1/1	1223	1244	.. 1	23 [.
mRNA_cap_enzyme	1/1	1004	1309	.. 1	504 []

	Score	E-value
pkinese	69.4	3.7e-18
pkinese	22.8	3.2e-05
pkinese	51.2	4.2e-13
pkinese	138.7	1.1e-37
Ribosomal_L23	5.0	3.9
mRNA_cap_enzyme	-181.3	9.6

Alignments of top-scoring domains:

Pkinese: domain 1 of 4, from 332 to 443: score 69.4, E = 3.7e-18

```

*->ilk..kesls..lrEiqilkrslsHpNIvrllg
      +++++ k++++++ +E 1 +lsHpN+vr+l
14790 332 KIDkcKKQIQgtETEFNSLVKLSHPNVVRYLA 363

vfed...tddhlylvmEymegGdLfdylrrngpls
+ +++++ + + E+ g +L+ +l+ gp++
14790 364 MNLKeqdDISVVDILVEHISGVSLAAHLSHSGPIP 398

ekeakkialQilrGleYLHsngivHRDLKpeNILl
+++++++ Q+l+Gl+YLHsn++vH L ++N+L+

```

FIG. 7I



34/53

14790 399 VHQLRRYTAQLLSGLDYLHSNSVVHKVLSASNVLV 433

dengtvKiaD<-* SEQ. ID NO: 8

d +gtvKi+D

14790 434 DAEGTVKITD 443

pkinese: domain 2 of 4, from 501 to 539: score
22.8, E = 3.2e-5

*->rlplpsncSeelkdLlkkcLnkDPskRpGsat

+ +ps ++ +++d+lkkc ++D ++R+ +

14790 501 PVTIPSDLPADFQDFLKKCVCLDDKERW---S 529

akeilnhpwf<-* SEQ. ID NO: 9

+++l+h ++

14790 530 PQQLLKHSFI 539

pkinese: domain 3 of 4, from 590 to 662: score
51.2, E = 4.2e-13

*->yelleklGeGsfgkVykakhk.tgkivAvKil

+e l+ lG+G+fG V k+++k +g+ +AvK +

14790 590 FEELQLLGKGAFGAVIKVQNKlDGCCYAVKRI 621

kkesls.....lrEiqilkrslsHpNIvrllgvfe

+ s++ ++ E+ +l rl+H+NIvr++ ++

14790 622 PINPASrqfrriKGEVTLLSRLHHENIVRYNAWI 656

O.G. FIG.	SUBCLASS
	CLASS
APPROVED BY	DRAFTSMAN



35/53

APPROVED	O.G. FIG.	SUBCLASS
	CLASS	
BY	DRAFTSMAN	

```
dtddhl<-* SEQ. ID NO: 10
++++++
14790 657 ERHERP      662

pkinase: domain 4 of 4, from 797 to 1001: score
138.7, E = 1.1e=37

* <-hlylvmEymegGdLfdylrrngplsekeakki
  +ly+ mEy+e+  L+d + + + + + ++
14790 797      YLYIQMEYCEKSTLRDTIDQGLYRDTVRLWRL      828

alQilrGleYLHsngivHRDLKpeNILldengtV
+++il Gl+Y+H++g +HRDLKp NI+ld++ +vK
14790 829 FREILDGLAYIHEKGMiHRDLKPVNiFLDSDDHVK      863

iaDFGLArll.....ekl
i+DFGLA++      + ++++++++ ++++++l
14790 864 IGDFGLATDHlafsadskqddqtdliksdpsGHL      898

ttfvGTpwYmmAPEvileg...rgysskvDvWSlG
t+ vGT  Y+ +PEv  +g++++ Y kvD  SlG
14790 899 TGMVGTALYV-SPEV--QGstkSAYNQKVDLFSLG      930

viLyElltggplfpgadlpaftggdevdqliifvl
+i++E+
14790 931 IIFFEMS----- 937

klPfsdelpktridpleelfrikkr....rlplps
  P                      e++f +++ ++++ +++p+
```

FIG. 7K



36/53

14790 938 YHPMV-----TASERIFVLNQLrdptSPKFPE 964

ncSee....lkdLlkkcLnkDPskRpGsatakeil
+ + ++ + k+++++Ln DP+kRp ta+e+l

14790 965 DFDDGehakQKSVISWLLNHDPakRP---TATELL 996

nhpwf<-* SEQ. ID NO: 11
++ +

14790 997 KSELL 1001

Ribosomal_L23: domain 1 of 1, from 1223 to 1244:
score 5.0, E = 3.9

->tdiikyPviTeKlamnlleepNk<-
↑SEQ. ID NO: 12
++ii y +TeKl++++e ++

14790 1223 VYIILYDAVTEKLTRREVEA-KF 1244

RNA_cap_enzyme: domain 1 of 1, from 1004 to 1309:
score -181.3, E = 9.6

*->nqtteRvyelhkiElfsvpelnGKKiglgil
q++e + e+ l

14790 1004 PQMEES-----ELHEV-----LHH 1017

kLpktdteslrtnVakllglamktktfPddegS
L+++d++ rtm a+ + P +

14790 1018 TLTNVDGKAYRTMMAQIFS----QRISPAIDYT 1046

qPVsferkdleesLkekdyfvceKTDGircshg
+ + + ++ vce ir+

14790 1047 YDSDILKGNFSIRTAKMQQHVCETI--IRI--- 1074

FNRTGFLIAaLlFlvehpgleeaiSHiLSgef.
+ +h + +++

14790 1075 -----FKRHGAVQL-----CTp 1086

.lidReknyYKQDYIDllpkrlfPrekdktkak
l+ R +Y ++ ++ lf

14790 1087 lLLPRNRQIY-----EHNEAALFMD----- 1106

APPROVED BY DRAFTSMAN	O.G. FIG.	SUBCLASS
	CLASS	

FIG. 7L



37/53

APPROVED BY DRAFTSMAN	O.G. FIG.
	CLASS SUBCLASS

elptyhrgtllDGEIvidinriaveqkTlrYvv
+ + l++ l d ++ rYv
14790 1107 -----HSGMLVM-LPFD----LRIPF-ARYVA 1127

FDalaisGqtviqrd.lskrLgdefikavkKpf
++ ++ + i+r + L+ f
14790 1128 RNNILNLKRYCIERVfRPRKLDLDR-F-----HP 1153

defkkvmpdakilnqqkYNFpfkiglkhmslsy
+e+ d+++++ f++ + ++y
14790 1154 KELLECAFDIVTSTT---NSFLPTAEIIYTIY 1182

gqlkllkaeskmviskadampkllHindGllft
++ + a +
14790 1183 EIIQEFPA-----L----- 1191

cvrdtpyieGeiLVEPGNSYlDfnLlKWKpkee
++++i+ + LlK
14790 1192 QE-RNYSIYL-----NHTMLLKA----- 1208

nTvDFelilefeevndPeldekdgfslyLdYda
+l + +e++
14790 1209 -----ILLHCG-----IPEDK----- 1219

mpGELfkfslgVWqgGfnkrFevihtdqiffrv
+ ++++ + ++++++ v
14790 1220 -----LSQVYIILY-DAVTEKLTRRE-----V 1240

afqklgRlkhefaelsVsdkdwyklkaleqpld
++ f +ls+s +l
14790 1241 EAK-----FCNLSLSSNSLCRLY----- 1258

GrIVEcrladieilIFQegrWeylrfRdDKqqa
+++E ++ + + ++ + q+
14790 1259 -KFIEQKGDLDQD-----LMPTINSLIKQK 1281

lKtgGYsgNhistvekvllsikDgvsieEeLlkl
tg +++v kD +Ee+ l
14790 1282 --TG-----IAQLVKYGLKD---LEEUVGL 1301

FIG. 7M



38/53

fpGmyFAGAktlikr<-* SEQ. ID NO: 13

k +ik

14790 1302 LK-----KLGIKL 1309

//

Searching for complete domains in SMART
hrmpfam - search a single seq against HMM database
HMMER 2.1.1 (Dec 1998)
Copyright (c) 1992-1998 Washington University
School of Medicine
HMMER is freely distributed under the GNU General
Public License (GPL).

HMM file: /ddm/robison/smart/smart/smart.all.hmms
Sequence file: /prod/ddm/wspace/orfanel/oa-
script.16895.seq

Query: 14790

Scores for sequence family classification (score
includes all domains):

Model	Description	Score	E-value	N
serkin_6		184.2	2.2e-51	2
tyrkin_6		-40.3	2e-09	1

Parsed for domains:

Model	Domain	seq-f	seq-t	hmm-f	hmm-t
serkin_6	1/2	286	539..	1	231 []
serkin_6	2/2	590	1001..	1	231 []
tyrkin_6	1/1	590	1001..	1	280 []

Model	score	E-value
serkin_6	17.2	2.2e-11
serkin_6	166.8	3.5e-46
tyrkin_6	-40.3	2e-09



39/53

Alignments of top-scoring domains:

serkin_6: domain 1 of 2, from 286 to 539: score
17.2, $E = 2.2e-11$

O.G. FIG.	SUBCLASS
	CLASS
APPROVED BY	DRAFTSMAN

```
*->YellkklGkGaFGkVylardkktgrlvAiKvi
      ++lGk      Vy a ++ tg v      +
14790 286 IGSDEQLGK----LVYNALETATGGFV---LL 310

      k.....erilr
      +   + +++ ++ ++++++ ++ +++ + ++
14790 311 YewvlqwqkkmgpfltsqekekidkckkqiQGTET 345

      EikiLkk.dHPNIVkLydvfed.....dklylVmE
      E + L k  HPN+V+++ +   ++++++ + + E
14790 346 EFNSLVKlSHPNVVRYLAMNLKeqddsIVVDILVE 380

      yceGdlGdLfdllkkrgrrglrkvlsE.earfyfr
      +++G   +L  +l + g+   ++ ++ r+y++
14790 381 HISG--VSLAAHLSHSGP-----IPVhQLRRYTA 407

      QilsaLeYLHsqqIiHRDLKPeNiLLds..hvKla
      Q+ls+L+YLHs+ ++H  L  +N+L+d +++vK++
14790 408 QLLSGLDYLSNSVVKVLSASNVLVDAegTVKIT 442

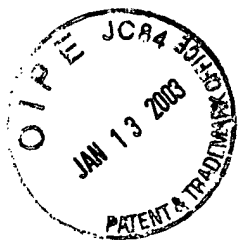
      DFGlArql.....ttfvGTpeYmAPEvl..
      D  ++++l +  ++  + t v      +++
14790 443 DYSISKRLadickedvfeQTRV-----RFsd 468

      .....gYgkpavDiWSlGcilyElltGkpPFpqlld
      +  + + gk + D+W lG +l  l  G+
14790 469 nalpyKTGK-KGDVWRLGLLLLSLSQGE-----C 497

      lifkkig....SpeakdLikklLvkdPekRlta.e
      +  +  +++ + ++ d++kk+ + d ++R++ ++
14790 498 GEYPVTIpsdlPADFQDFLKKCVCLDDKERWSPqQ 532

      aLedeldikaHPFF<-* SEQ. ID NO: 14
      +L+      H f+
14790 533 LLK-----HSFI 539
```

FIG. 70



40/53

serkin_6: domain 2 of 2, from 590 to 1001: score
166.8, E = 3.5e-46

APPROVED BY DRAFTSMAN	O.G. FIG.	SUBCLASS
	CLASS	

```
*->YellkkkLGkGaFGkVylardkktgrlvAiKvi
      +e l++lGkGaFG V ++++k +g +A+K+i
14790 590 FEELQLLGKGAFGAVIKVQNKLDGCCYAVKRI 621

      k.....
      + ++ +++ ++ +++ + ++ ++++ + +
14790 622 Pinpasrqfrrikgevtllsrlhhenivryynawi 656

      .....
      ++++++ ++++++ ++++++ ++++++ ++++++
14790 657 erherpagpgtpppdsgplakddraargqpasdtd 691

      .....
      + ++ + +++ +++ + ++++++ + + +
14790 692 gldsveaaapppilsssviewstsgersasarfpat 726

      .....
      + ++ + +++ +++ + ++++++ + + +
14790 727 gpgssddedddedehggvfsqsflpasdsesdiif 761

      rilrEikiLkk...dHPNIVkLydvfed.....d
      + + E + +++++d+ +++++ + +++ +
14790 762 DNEDENSKSQNqdeDCNEKNGCHESEPSvtteavH 796

      klylVmEyceGdlGdLfdllkkrgrrglrkvlse.
      +ly+ mEyce +L+d +++ + +
14790 797 HLYIQMEYCEK--STLRDTIDQGLY-----RDTv 823

      earfyfrQilsaLeYLHsqqIiHRDLKPeNiLLds
      + +++fr+il++L+Y+H++g iHRDLKP Ni+Lds
14790 824 RLWRLFREILDGLAYIHEKGMiHRDLKPVNiFLDS 858

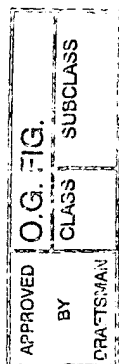
      ..hvKlaDFGLArql.....
      ++hvK++DFGLA+ + ++++++ +
14790 859 ddHVKIGDFGLATDHlafsadskqddqtgdlikd 893

      .....ttfvGTpeYmAPEvl.....gYgkpavDiW
```

FIG. 7P



41/53



14790 894 +++++ t +vGT Y++PEv +++++gY+ +vD
psghlTGMVGTALYVSPEVQgstksAYNQ-KVDLF 927

14790 928 SlGcilyElltGkpPFp..qldlifkkg.....
SlG+i++E+ + p ++ ++++++
SLGIIFFEMSY-HPMVTasERIFVLNQLRdptspk 961

14790 962SpeakdLikkllvkdPekRlta.eaL
+++ +++++ k+ i+ lL+ dP+kR+ta+e+L
fpedfddgeHAKQKSVISWLLNHDPKRPTAtELL 996

14790 997 edeldikaHPff<-* SEQ. ID NO: 15
+ + +
K-----SELL 1001

tyrkin_6: domain 1 of 1, from 590 to 1001: score -
40.3, E = 2e-09

14790 590 *->ltlgkklGeGaGFeVykgtlk...ieVAVKtL
+ ++ LG GaFG V k + k ++ AVK +
FEELQLLGKGAFGAVIKVQNKldgCCYAVKRI 621

14790 622 keda.....keeFlrEakiMkklGgkHpNiVklLgV
+ +++ + + E +++++l +H+NiV+ + +
PINPasrqFRRIKGEVTLLSRL--HHENIVRYNA 654

14790 655 cteegrrFmevePlmivmEymegGdLldyLrknrp
+e
WIE-----RHERP 662

14790 663 k.....
+++++ +++++ +++++ +++++ ++ +
Agpgtpppdsgplakddraargqpasdtdgldsve 697

14790 698
+++ +++++ + + ++++++ + + + ++++++
aaapppilsssvewstsgersasarfpatgpgssd 732

.....
+++++ +++++ +++++ +++++

FIG. 7Q



42/53

APPROVED	C.C. F. J.	SUBCLASS
BY	CLASS	
CRAFTSMAN		

14790 733 dedddedehggvfsqsflpasdsesdiifdneden 767
.....
+++++++ +++++ +++++ +++ + +
14790 768 sksqnqdedcnekngcheseptvtteavhylyiqm 802
.....lslsdLlsfAlQIAkGMe
+ +++++ +++ +++ + L I +G +
14790 803 eycekstlrdtidqglyRDTVRLWRLFREILDGLA 837
YLesknfvHRDLAARNcLvgenkvvKIsDFGLsRd
Y+++k+ +HRDL N+ +++++ +vKI+DFGL+ d
14790 838 YIHEKGMiHRDLKPVNIFLDSDDHVKIGDFGLATD 872
lyddDkkG.....eskdyYrkkggkkggkttlPir.
+ ++ +++++ +d +++ +++ +++
14790 873 HLAf-SADskqddQTGDLIKSDPSGHLTGMVGTA1 906
WmAPESl..kdgkFtskSDVWSFGVlLWEiftlGe
+PE+ ++ ++ ++ k D +S G+ +E+
14790 907 YVSPEVQgsTKSAYNQKVDLFSLGIIFFEM----- 936
qPYpgeiqqfmsnee...vleylkkGyRlpkPend
Y + +++++e+ vl++l++ ++ Pe+
14790 937 -SYHPM----VTASerifVLNQLRDPTSPKFPED- 965
lpiSs.vtCPdelYdlMlqCwaedPedRPtFsel.
++ + +++ +++++ ++dP++RPT +el
14790 966 ----FdDGEHAKQKSVISWLLNHDPakRPTATELl 996
.verl<-* SEQ. ID NO: 16
+ e+l
14790 997 kSELL 1001



43/53

ProDom Matches

O.G. FIG.	CLASS	SUBCLASS
	APPROVED BY	DRAFTSMAN

ProdomId	Start	End	Description	Score
View Prodom 40346	18	323	p99.2 (2)001712(1) // INITIATION FACTOR KINASE EUKARYOCTIC-2 ALPHA EIF-2ALPHA	295
View Prodom 137719	19	1504	p99.2(1)074297_NEUCR// CPC3 PROTEIN	87
View Prodom I	341	1017	p99.2(2773)CC2(14) KKIT (14) KPC(13)//KINASE PROTEIN TRANSFERASE ATP- BINDING SERIN/THREOINE- PROTEIN PHOSPHORYLATION RECEPTOR TYROSINE-PROTEIN PRECURSOR TRANSMEMBRANE	107
View Prodom 150228	1105	1206	p99.2(1)074297_NEUCR// CPC3 PROTEIN	90
View Prodom 2305	1106	1230	p99.2(23)SYH(12) // SYNTHETASE AMINOACYL-TRNA HISTIDYL-TRNA PROTEIN LIGASE ATP-BINDING BIOSYNTHESIS HISTIDINE-- TRNA HISRS KINASE	137
View Prodom 42726	1274	1648	p99.2(2)0017121(1) 061651 (1)//INITIATION FACTOR KINASE EUKARYOTIC EIF-2 ALPHA EIF-2ALPHA	221

FIG. 7S



44/53

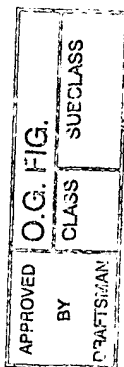
View Prodom 40346

>40346 p99.2(2) O01712(1) O61651(1) //INITIATION
FACTOR KINASE EUKARYOTIC EIF-2 ALPHA EIF-2ALPHA
Length = 296

Score = 295(108.9 bits), Expect = 1.7e-28, Sum
P(2) = 1.7e-28 Identities = 58/144 (40%),
Positives = 95/144 (65%)

Query:	18	SYPQRQDHELQALEAIYGADFQDLRPDACGPVKEP	52
		S+ +RQ EL+ +++I+G D +DLRP A + +P	
Sbjct:	9	SFRERQAQELEVIKSIFGCDVEDLRPQANPSLWKP	43
Query:	53	PEINLVLYP--QGLTGEEVYVKVDLRVKCPPTYPD	85
		+I + L P G E YV L V CP YP	
Sbjct:	44	TDIRIQLTPLRDSSNGLETYVCTKLHVTCPSKYPK	78
Query:	86	VVPEIELKNAKGLSNESVNLLKSRLEELAKKHCGE	120
		+ P+I L+ +KG+S++ + L+++L+ +++ GE	
Sbjct:	79	LPPKISLEESKGMSDQLLEALRNQLQAQSQELRGE	113

FIG. 7T





45/53

Query: 121 VMIFELAYHVQSFLSEHNKPPPKSFHEEMLERRAQ 154
VMI+ELA VQ+FL EHNKPP SF+++ML+ + +

Sbjct: 114 VMIYELAQTVQAFLLLEHNKPPKGSFYDQMLQDKQK 148

Query: 155 EEQQ 159
+Q+

Sbjct: 149 RDQE 152 SEQ. ID NO: 17

Score = 63 (27.2 bits), Expect = $1.7e-28$, Sum P(2) = $1.7e = 28$ Identities = 15/59 (25%), Positives = 30/59 (50%)

Query: 266 EILYFN-MGSPDQLMVHKGKCIQSDEQLGKLVYNA 299
E LYF+ MG + +G C+G ++ G + Y

Sbjct: 230 ETLYFHKMGR----QIQRGCCVGHSSQR-GCIAYTG 259

Query: 300 LETATGGFVLLYEWVLQWQKKMGP 323
++ G + + EW +++ + P

Sbjct: 260 IDMHCGQLLYITEW?IKYSQLEQP 283
↑SEQ. ID NO: 18

View Prodom 42726

>42726 p99.2(2) O01712(1) O61651(1) //INITIATION
FACTOR KINASE EUKARYOTIC EIF-2 ALPHA EIF-2ALPHA
Length =469

Score = 221 (82.9 bits), Expect = $7.5e-15$, P = $7.5e-15$ Identities = 92/341 (26%), Positives 160/341 (45%)

fig 7i

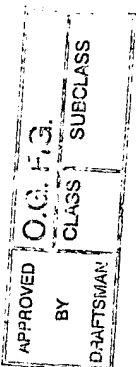
Query: 1274 INSLIKQKTGIAQLVKYGLKDLEEXXXXXXXXXX 1306
+ SL++ K A L + L++LE

Sbjct: 70 LKSLMRGKGEEAASLARGALRELETVVGLAYSLG 102

Query: 1307 XXXXXXXNLGLVYKVQQ--HNGIIFQFVAFIKR 1337
GL + + GI++Q A +K

Sbjct: 103 VKCPIHIWAGLPISFDRASNGGIVWQMTADLKP 135

FIG. 7U





46/53

APPROVED	BY	CLASS	SUBCLASS
O.G. FIG.			

Query: 1338 RQRAVPEILAAGGRYDLLIPQF-RGPQALGPVP 1369
+ P +LA G RYD ++ +F + Q P
Sbjct: 136 NRS GHPSVLAIGERYDSMLHEFQKQAQKFNPM 168

Query: 1370 TAIGV-----SIAIDKISAAVLNMEESVTISS 1396
A GV + ++DK+ AAV +E + +
Sbjct: 169 PARGVLSGAGLTFSLDKLVAAV-GVEYAKDCRA 200

Query: 1397 CDLLVSVSGQMSMSRAINLTQKL-WTAGITAEI 1428
D+ + G + + +L W+ GI I
Sbjct: 201 IDVGICVCGTRPPLKDVITYIMRLLWSVGIRCGI 233

Query: 1429 MYDWSQSQEELQEYCRHHEITYVALVSDKEGSH 1461
+ S+ +E Q+ R + +V LV++ GS
Sbjct: 234 VEAASELGDEAQLARLGAL-HVILVAEN-GS- 263

Query: 1462 VKVKSFEKERQTEKRVLETELV DHVLQKLRTKV 1494
++V+SFE+ER E+ + TELV+ + + LR+
Sbjct: 264 LRVRSEFERERFQERHLTRTELVEFIQKMLRS-- 294

Query: 1495 TDERNGREASDNLAVQNLKGSFSNAS----- 1520
D NG DN + + GS N S
Sbjct: 295 -DGLNGTTV-DNFHLSALGSGDNRSSGGKERE 325

Query: 1521 ----GLF-EIGHATV-----VPIVSV--LAPE 1540
GL AT+ +P + V L +
Sbjct: 326 RGENGLSTSASNATIKNNYSQLPNLQVTF LTHD 358

Query: 1541 KLSASTRRRYETQVQTRLQTS LANLHQKSSEIE 1573
K +A+ +RR E QV ++ ++L+ +K + +
Sbjct: 359 KPTANYKRRL ENQVAQQMSSTLSQFLKKETFV- 390

Query: 1574 ILAVDLPKETI 1584
+L V+LP +
Sbjct: 391 VLVVELPPAVV 401 SEQ. ID NO: 19

Score = 150 (57.9 bits), Except = 4.2e-07, P = 4.2e-07
Identities = 66/300 (22%), Positives = 137/300 (45%)



47/53

APPROVED	O.G. FIG.	
	CLASS	SUBCLASS
BY	DRAFTSMAN	

Query: 1368 VPTAIGVSI AIDKISA AAVLNMEESVTISSCDLL 1400
V + G++ ++DK+ AAV +E + + D+

Sbjct: 173 VLSGAGLTFS LDKLVAAV-GVEYAKDCRAIDVG 204

Query: 1401 VVSVGQMSMSRAINLTQKL-WTAGITAEIMYDW 1432
+ G + + +L W+ GI I+

Sbjct: 205 ICVCGTRPPLKDV TYIMRLLWSVGIRCGIVEAA 237

Query: 1433 SQSQEELQEYCRHHEITYVALVSDKEGSHVKVK 1465
S+ +E Q+ R + +V LV++ V+

Sbjct: 238 SELGDEA QDLARLGAL-HVILVAENGLRVR SF 269

Query: 1466 SFEK--ERQTEKRVLETETELVDHVLQK--LRTKV 1494
E+ ER + L E + +L+ L

Sbjct: 270 ERERFQERHLTRTEL-VEFIQKMLRSDGLNGTT 301

Query: 1495 TDERNGREA---SDNLAV-----QNLKGSFSN 1518
D + A DN + + G ++

Sbjct: 302 VDNF SHLSALGSGDNRSSGGKERERGENGLSTS 334

Query: 1519 ASGLFEIHGATVVPIVSV--LAPEKLSASTRRR 1549
AS + + +P + V L +K +A+ +RR

Sbjct: 335 ASNATIKNNYSQLPNLQVTFLTHDKPTANYKRR 367

Query: 1550 YETQVQTRLQTS LANLHQKSSEIEILAVDLPKE 1582
E QV ++ ++L+ +K + + +L V+LP

Sbjct: 368 LENQVAQQMSSTLSQFLKKETFV-VLVVELPPA 399

Query: 1583 TILQFLSL--EWD ADEQAFNTTVKQLLSRLPK- 1612
+ + + ++ + ++ R K

Sbjct: 400 VVNAIVGAINPREIRKRETEPEINYVIERFSKY 432

Query: 1613 QRYLKLVCDEIYNIKVEKKVSVLFLYSYRDDYY 1645
+RY+ + +E+ + + K ++ LYS D YY

Sbjct: 433 KRYISEINEEVVDYLSDAKTPIVALYSISDSYY 465

Query: 1646 RIL 1648
R++

Sbjct: 466 RVI 468 SEQ. ID NO: 20



48/53

View Prodom 2305

>2305 p99.2 (23) SYH(12) // SYNTHETASE AMINOACYL-
TRNA HISTIDYL-TRNA PROTEIN LIGASE ATP-BINDING
BIOSYNTHESIS HISTIDINE--TRNA HISRS KINASE Length =
145

Score = 137 (53.3 bits), Expect = 3.9e-08, P =
3.9e-08 Identities = 41/131 (31%), Positives =
62/131 (47%)

Query: 1106 DHSGMLVMLPFDLRIPFARYVARNNI----LNL 1134
D G L+ L +DL +PFARYVA N + L L
Sbjct: 18 DQGCELLSLRYDLTVPFARYVAMNLLKVTNLPL 50

Query: 1135 KRYCIERVFRPRK--LDRFHPKELLECAFDIVT 1165
KRY I +V+R + + R +E +C FDI+
Sbjct: 51 KRYHIAKVYRRDRPAMTRGRYREFYQCDFDII- 82

Query: 1166 STTNSFLPXXXXXXXXXXXXXQEFPALQERNYSI 1198
++ P + + N+ I
Sbjct: 83 GEYDTMAPDAEILKILTEILSQLGIRELGNFKI 115

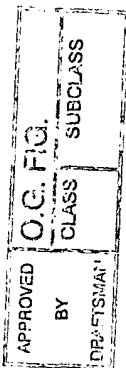
Query: 1199 YLNHTMLLKAILLHCGIPEDKLSQVYIILYDA 1230
+NH +L ++L P K Q Y+ Y A
Sbjct: 116 KINHRGILDSLLQ----PWPKTLQEYLTQYKA 143
↑SEQ. ID NO: 21

View Prodom 1

>1 p99.2 (2773) CC2(14) KKIT(14) KPC1(13) //KINASE
PROTEIN TRANSFERASE ATP-BINDING SERINE/THERONINE-
PROTEIN PHOSPHORYLATION RECEPTOR TYROSINE-PROTEIN
PRECURSOR TRANSMEMBRANE Length = 431

Score = 107 (42.7 bits), Expect = 0.0032, Sum P(2)
= 0.0032 Identities = 36/106 (33%), Positives =
50/106 (47%)

FIG. 7X





49/53

Query: 848 DLKPVNIFLDSDDH-----VK-IGDFGLATDHLA 875
DLKP NI LD + H +K I DFGLA +

Sbjct: 220 DLKPENILLDEESHENTPNMIKLIADFLAKE--I 253

Query: 876 FSADSKQDDQTGDLIKSDPSGHLTGMVGTALYVS- 909
+S+ S ++ + + M GT YVS

Sbjct: 254 YSSSSTYEEMSSSQAVFGSHQTTSTMCGTPYYVSM 288

Query: 910 ----PEVQGSTKSA-----YNQKVDLFSLGIIFFE 935
PE SA Y+ K D++S G+I +E

Sbjct: 289 KSMAPPEYMAPESSATNYQKYSTKSDVWSFGVILYE 323

Query: 936 M 936
M

Sbjct: 323 M 323 SEQ.ID NO: 22

Score = 105 (42.0 bits), Expect = 3.5e-05 Sum
P(3)=3.5e-05 Identities = 35/102 (34%), Positives
= 48/102 (47%)

Query: 824 RLWRLFREILDGLAYIHEK-----GMIHR----DL 849
+L +I GL Y+H K G+IHR DL

Sbjct: 187 QLMHYVHQIAKGLEYLHSKNQKHQGIHRRAKKVDL 221

Query: 850 KPVNIFLDSDDH-----VK-IGDFGLATDHLAFS 877
KP NI LD + H +K I DFGLA + +S

Sbjct: 222 KPENILLDEESHENTPNMIKLIADFLAKE--IYS 254

Query: 878 ADSKQDDQTGDLIKSDPSGHLTGMVGTALYVS 909
+ S ++ + + M GT YVS

Sbjct: 255 SSSTYEEMSSSQAVFGSHQTTSTMCGTPYYVS 286
↑SEQ.ID NO: 23

Score = 74 (31.1 bits), Expect = 0.91, Sum P(2) =
0.60 Identities = 30/127 (23%), Positives = 60/127
(47%)

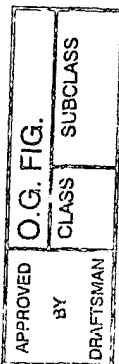
Query: 341 QGTETEFNSLVKLSHPNVVRYLAMNLKEQDDSIIVV 375
+G+ E+ + + ++ L++ + +++

Sbjct: 134 EGSLVEYMEYMSGGSEDYMKKLSLETVMKIAMMIL 168

FIG. 7Y



50/53



Query: 376 DIL-VEHISGVSLAAHLSHSGPIPVHQLRRYTAQL 409
+ + H+S S + LSHS QL Y Q+

Sbjct: 169 QFMQIMHMSSESES--LSHS-----QLMHYVHQI 195

Query: 410 LSGLDYLHSNS-----VVHKV----LSASNVLVDA 435
GL+YLHS + ++H+ L N+L+D

Sbjct: 196 AKGLEYLHSHKNQKHQGIHRAKKVDLKPENILLDE 230

Query: 436 EG-----TVK-ITDYSISKRL 450
E +K I D+ ++K +

Sbjct: 231 ESHENTPNMIKLIADFGLAKEI 252
↑SEQ.ID NO: 24

Score = 65 (27.9 bits), Expect = 3.5e-05, Sum P(3) = 3.5e-05
Identities = 29/124 (23%), Positives = 50/124 (40%)

Query: 907 YVSPEVQGSTKSAYNQKVDLFSLGIIFFEM-SYHP 940
Y++PE + Y+ K D++S G+I +EM + P

Sbjct: 294 YMAPESSATNYQKYSTKSDVWSFGVILYEMLTGKP 328

Query: 941 MVTASERIFVLNQLRDPTSPKFPEDFDDGEHAKQK 975
E +++ S K E + G +

Sbjct: 329 PFFPGES--EVSEEEPYQSMKNMEVLEMGPETIQ 361

Query: 976 SVISWLLNHDPKRPT-----ATELLKS 998
V+S ++ + P A +LLK

Sbjct: 362 KVMSKIVEKKGERMPQPSSSNCPEVSQEAKDLLKK 396

Query: 999 ELLPPPQMESELHEVLHH 1017
L P+ E+L H

Sbjct: 397 CLQKDPEKRRPTFEEILQH 415 SEQ.ID NO: 25

Score = 55 (24.4 bits), Expect 3.5e-05, Sum P(3) = 3.5e-05
Identities = 12/23 (52%), Positives 18/23 (78%)

Query: 589 EFEEELQ-LLGKGAFGAVIKVQNK 610
++E L+ LLGKG+FG V K ++K

Sbjct: 33 QYELLKKLLGKGSFGKVYKAKHK 55
↑SEQ.ID NO: 26

FIG. 7Z



51/53

Score = 49 (22.3 bits), Expect 34., Sum P(2) = 1.0
Identities = 12/39 (30%), Positives = 24/39 (61%)

Query: 507 DLPADFQDFLKKCVCLD-DKERWSPQQLLKHS-FI 539
++ + +D LKKC+ D +K R + +++L+H F+
Sbjct: 385 EVSQEAKDLLKKCLQKDPEKRRPTFEEILQHPWFL 419

Query: 540 NPQP 543
P

Sbjct: 420 MRNP 423 SEQ.ID NO: 27

Score = 40 (19.1 bits), Expect = 0.0010, Sum P(3)
= 0.0010 Identities = 8/11 (72%), Positives = 9/11
(81%)

Query: 596 LGKGAFGAVIK 606
LG G+FGAV K

Sbjct: 2 LGTGSFGAVYK 12 SEQ.ID NO: 28

View Prodom 150228

>150228 p99.2 (1) O74297_NEUCR //CPC3 PROTEIN
Length = 108

Score = 90 (36.7 bits), Expect = 0.0039, P =
0.0039 Identities = 32/105 (30%), Positives =
43/105 (40%)

Query: 1105 MDHSGMLVMLPFDLRIPFARYVAR--NNILN 1133
+D +G ++ LPFDL + AR +AR N+ +
Sbjct: 3 LDQNGTVLQLPFDLMMGHARSLARITNSPVV 33

Query: 1134 LKRYCIERVFRPRKLD RFHPKELLECAFDIVT 1165
K Y +FR R P E FDI VT
Sbjct: 35 QKSYSFGNIFRDRH-GGGQPDVYGEVDFDIVP 65

Query: 1166 STTNSF-LPXXXXXXXXXXXXXQEF PALQERNY 1196

FIG. 7A1



52/53

S L FP +
Sbjct: 66 SDALDLALKEAEVIKVLDEIATAFPTVSSTPI 97

Query: 1197 SIYLNHTMLL 1206

L H+ LL

Sbjct: 97 CFQLGHSDLL 106 SEQ.ID NO: 29

View Prodom 13771

>137719 p99.2(1) 074297_NUCR //CPC3 PROTEIN Length
= 304

Score = 87 (35.7 bits), Expect 0.61, Sum P(2) =
(0.45)Identities=41/146 (26%), Positives = 64/156
(41%)

Query: 19 YPQRQDHELQALEAIYGADFQDLRPDACGPVKEPP 53
Y + Q+ E+ L+AIYG DF K P

Sbjct: 44 YQEVQSESEVMVLQAIYGEDFTQHEAAHGAWQKSEP 78

Query: 54 EINLVLYPQGLTGEEVYVKVDLRVKCPPTYPDVVP 88
++ + P + +E+ V L V TYP P

Sbjct: 79 RFDIKIKPS--SDQEL--SVTLGVVMVATYPKTPP 109

Query: 89 EIELKNAKGLSNESVNLLKSRLEELAKK---HCGE 120
+ +K+ L ES + E K +

Sbjct: 110 LLTIKDDHSL-RESTKFQKFKVETQPKIYAQAEQ 143

Query: 121 VMIFELAYHVQSFLSE--HNK-----PP---PKS 144
MI ++ ++ L E K P ++

Sbjct: 144 EMIDQIVEGIRDILEEAAQKKVQGLEIPSLEEERA 178

Query: 145 FHEEMLERRAQEEQQR 160
HE L R AQ E++R

Sbjct: 179 AHEAELARLAQSEKER 194 SEQ.ID NO: 30

Score = 49 (22.3 bits), Expect =0.61, Sum P(2) =
0.45 Identities = 13/48 (27%), Positives = 27/48
(56%)

FIG. 7A2



53/53

Query: 1458 EGS HVK V K S F E K E R Q T E K R V L E T E L V D H V L 1487
E ++ E K E R + K ++ E ++ + V L

Sbjct: 181 E A E L A R L A Q S E K E R E E R K K L E E S K E E E R V L 210

Query: 1488 Q K - L R T K V T D E R N G R E A S 1504
+ L + ++ + R N + S

Sbjct: 211 E D M L Q E E L K R Q R N K A K E S 228 SEQ.ID NO: 31

Score = 49 (22.3 bits), Expect = 0.61, Sum P(2) =
0.45 Identities = 13/48 (27%), Positives = 27/48
(56%)

Query: 238 K H R A N S S G R S R R E R Q Y S V C N S E D S P G S C - E I L 268
+++ A S + R Q S + P G E L

Sbjct: 222 R N K A K E S R K K N R S H Q L S P D R A P Q D P G E T D E T L 253

Query: 269 Y F N M G S P D Q L M V H K G K C I G S D E Q L G K L V Y 297
F + P ++ G + + G K V +

Sbjct: 254 M F D Q -- P C K I T D G S G N A L F F Q T V I G K T V F 280
SEQ.ID NO: 32

Score = 47 (21.6 bits), Expect = 0.96, Sum P(2) =
0.62 Identities = 19/83 (22%), Positives = 33/83
(39%)

Query: 750 L P A S D S E S D I I F D N E D E N S K S Q - N Q D E D C N E K N G C 783
L S E ++ D E K Q N + ++ + K N

Sbjct: 200 L E E S K E E E R V L E D M L Q E E L K Q R Q R N Y A K E S R K K N R 234

Query: 784 H E S E P S V T T E A V H Y L Y I Q M E Y C E K S T L R D T I D Q G L 818
H + P + + + + D L

Sbjct: 235 H Q L S P D R A P Q D P G E T D E T L M F D Q P C K I T D G S G N A L 269

Query: 819 Y R D T V R L W R L F R E 831
+ T V + F R E

Sbjct: 270 F F Q T V I G K T V F R E 282 SEQ.ID NO: 33

APPROVED BY DRAFTSMAN	O.G. FIG.	SUBCLASS
	CLASS	